

Use of Automatically Collected Data to Improve Transit Line Performance

by

Erik Sheridan Wile

Sc.B. Computer Science, Brown University (1997)

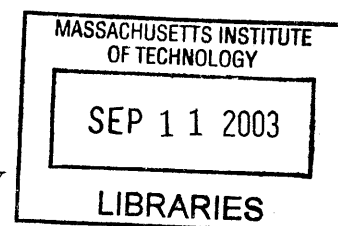
Submitted to the Department of Civil and Environmental Engineering
in partial fulfillment of the requirements for the degree of

Master of Science in Transportation

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

September 2003



© Massachusetts Institute of Technology 2003. All rights reserved.

Author
Department of Civil and Environmental Engineering
June 13, 2003

Certified by
Nigel H. M. Wilson
Professor of Civil and Environmental Engineering
Thesis Supervisor

Certified by
Haris Koutsopoulos
Associate Professor of Civil and Environmental Engineering
Northeastern University
Thesis Supervisor

Accepted by
Oral Buyukozturk
Chairman, Department Committee on Graduate Students

BARKER

Use of Automatically Collected Data to Improve Transit Line Performance

by

Erik Sheridan Wile

Submitted to the Department of Civil and Environmental Engineering
on June 13, 2003, in partial fulfillment of the
requirements for the degree of
Master of Science in Transportation

Abstract

This thesis investigates the potential for transit agencies to more fully exploit the data collected by a range of presently operational information and control systems. These systems – such as Automated Vehicle Location systems and Automated Fare Collection systems – are often purchased specifically to improve one aspect of agency performance – such as safety or revenue collection. They may alternatively be viewed as sources of information on vehicles and passengers, the building-blocks of an agency-wide effort to improve information on the state of the transit system, facilitating a deeper understanding of system dynamics than is allowed by manual data collection methods alone.

This thesis presents the case for the latter approach through the development of a framework for the appraisal and fulfillment of a range of agency functional needs using data from installed and proposed data collection systems. For each such need – including Service Management, Passenger Information, Performance Auditing, and eight others – applications which use automatically collected data are proposed. After an assessment of the relative benefit difficulty and benefit of each application, one or two are selected for implementation.

It then applies this framework to two case-study agencies: the Chicago Transit Authority [CTA] and Boston's Massachusetts Bay Transit Authority [MBTA], including implementation of three prototype applications which more fully utilize available agency data. The example applications are: 1. A system to provide web-accessible vehicle position reports using data from an on-vehicle emergency communications system installed on CTA buses. 2. Analysis of vehicle position over a long period of time to provide insight into line characteristics. 3. An application combining train position information and historical demand data to provide management daily passenger-centric line performance summaries. Each of these applications provides a number of demonstrable agency benefits as well as invaluable lessons for future application development within similar agencies. The benefits identified makes the case that there are indeed advantages to a holistic approach to use of automatically collected data in transit agencies.

Thesis Supervisor: Nigel H. M. Wilson

Title: Professor of Civil and Environmental Engineering

Thesis Supervisor: Haris Koutsopoulos

Title: Associate Professor of Civil and Environmental Engineering
Northeastern University

Acknowledgments

Thank you to Nigel Wilson for all the guidance, advice, compliments, criticism, encouragement and tolerance throughout this remarkable learning process. I am a much better person for having known him.

Thanks also to Haris Koutsopoulos for volumes of advice, largely orthogonal to Professor Wilson's, which consistently helped make this thesis something of which I can be proud.

Thanks to those who I have helped me at the CTA, especially Tom Pleuger for the CTA system information and for the use of the spare development server resources. Thanks to Angela Moore and Daniel Shurz and all the other users of the passenger information website for the encouragement and support. A second thank you to Angela Moore for a collection of high-quality data from CTA buses using floppy disks. Thanks also to the President of the CTA, Frank Kreusi, for supporting this research and being interested in the results. Thanks to Henry Marr at the MBTA for tending the automatic database dumps and Jeffrey Parker at the MBTA for helping to improve the usefulness of performance summaries.

Thanks to everyone in our research group, especially Ginny Siggia for making sure things got done, Joseph Coughlin for being a devil or devil's advocate [its unclear which], as well as Mikel Murga, Peter Furth, Ken Krukemeyer and Fred Salvucci for their own special variety of inspiration.

A most heartfelt thank you to my parents, Linn and David Wile, for decades of relentless, crucial, unconditional love and high expectations.

Thanks also to my friends who variously kept me focused, drove me crazy and kept me sane. Thanks to Ani DiFranco and Anne Heaton and the occasional Charles River Night Heron for reminding me why I get out of bed each day.

Contents

1	Introduction	15
1.1	Objectives	15
1.2	Motivation	16
1.3	Approach	18
1.4	Thesis Organization	19
2	The State of Practice	21
2.1	Agency Needs	22
2.1.1	Security / Safety	22
2.1.2	Service Management	23
2.1.3	Passenger Information	24
2.1.4	Performance Monitoring	25
2.1.5	Maintenance	25
2.1.6	Marketing	26
2.1.7	Operations Planning	26
2.1.8	Performance Auditing	27
2.1.9	Scheduling	29
2.1.10	Route Planning	30
2.1.11	Budgeting	30
2.2	User Data Needs	31
2.3	Automated Data Collection Systems	32
2.3.1	Security / Emergency Location	32
2.3.2	Automated Vehicle Location	33

2.3.3	Automatic Vehicle Identification	33
2.3.4	Automated Train Supervision and Control	34
2.3.5	Trip Time Analyzers	34
2.3.6	Automated Passenger Counters	35
2.3.7	Stop Announcement Systems	36
2.3.8	Electronic Fareboxes	36
2.3.9	System Integration	36
2.4	Technology	37
2.4.1	Location Technology	37
2.4.2	Communications Technology	40
2.5	Interface Standardization	41
2.6	Processing of Automatically Collected Data	42
2.6.1	Vendor Software	42
2.6.2	Specialized Analysis	43
2.6.3	Third-Party Tools	45
2.7	Barriers to Use	46
2.7.1	Data Formatting and Archival	47
2.7.2	Programmer Availability / Organizational Issues	47
2.7.3	Historical Lack of Standards	48
2.7.4	Security Through Obscurity	48
2.7.5	System Reliability	48
2.7.6	International GPS Skepticism	49
3	Method and Approach	51
3.1	Agency Data Sources	52
3.2	Data Source Utilization	52
3.3	Agency Needs and Opportunities	53
3.4	Intermediate Data Products	54
3.4.1	Application Challenges	56
3.5	Application Selection	57

4	Case 1: CTA Bus	61
4.1	Agency Appraisal	63
4.1.1	Agency Data Sources	63
4.1.2	Data Source Utilization	66
4.1.3	Agency Needs and Opportunities	68
4.1.4	Intermediate Data Products	79
4.1.5	Application Challenges	89
4.1.6	Application Selection	98
4.2	Agency Benefits	100
4.2.1	Real-time Information Tool	101
4.2.2	Post-Processing Data Analysis	106
4.2.3	Summary	119
4.3	Implementation Barriers and Solutions	120
4.3.1	Define Consistent Interfaces	120
4.3.2	Working With Available Resources	124
4.3.3	Working Within Collection System Limitations	127
5	MBTA Application	129
5.1	Agency Appraisal	129
5.1.1	Agency Data Sources	130
5.1.2	Data Source Utilization	133
5.1.3	Agency Needs and Opportunities	133
5.1.4	Intermediate Data Products	143
5.1.5	Application Challenges	152
5.1.6	Application Selection	161
5.2	Agency Benefits	162
5.2.1	Service Quality Auditing Application	164
5.2.2	Summary	177
5.3	Implementation Barriers and Solutions	177
5.3.1	Translate to Machine-Usable Formats	177

5.3.2	Build off standard, flexible, data processing tools	178
5.3.3	Provide Data Clearinghouse	179
6	Summary and Conclusions	183
6.1	Conclusions	184
6.1.1	Productive Framework	185
6.1.2	Significant Benefits of Using Data	185
6.1.3	Data Repository	186
6.1.4	Hidden Costs of Software Packages	192
6.1.5	Data Quality	193
6.2	Future Work	194
6.2.1	CTA Applications	194
6.2.2	MBTA Applications	197
6.2.3	Future Research	198
A	Example Database Usage	201
A.1	Vehicle Polling	201
A.2	Frequent Database Position Update	202
A.3	Example Position Database Use	204

List of Figures

3-1	Mapping from Data Collection System to Agency Use	53
3-2	Summary of Agency Data Needs	55
4-1	CTA System Map	62
4-2	Case 1: CTA Bus Use of Agency Technology	67
4-3	Position History Website	103
4-4	Position History Website	104
4-5	Trip Time on Route 77	107
4-6	Weekday Trip Time on Route 77	108
4-7	Operator Speed Variability on Westbound Route 77	109
4-8	Saturday Headway Degradation	110
4-9	Example Route 77 Eastbound Bus Data / 4 Minutes Polling Cycle . .	112
4-10	Example Route 77 Eastbound Bus Data / 70 Second Polling Cycle .	113
4-11	Example Route 77 Eastbound Bus Data / 4 Second Polling Cycle . .	114
4-12	Detail from Figure 4-11 with Door Open Tick Marks	115
4-13	Bus Communications Data Flow Diagram	117
5-1	MBTA System Map	130
5-2	MBTA Use of Agency Technology	133
5-3	Schematic of MBTA Red Line.	163
5-4	Red Line Exceptions excerpt for 4/1/02	166
5-5	Average Wait excerpt for 4/1/02	166
5-6	Wait-Time Percentages excerpt for 4/1/02	167
5-7	Vehicle Headway Percentages excerpt for 4/1/02	168

5-8	Example half day, one direction plot	169
5-9	Train trajectory and manually collected confirmation.	171
5-10	Average Passenger Wait Times (minutes) in the AM Peak.	171
5-11	1/10/02 Blue Line Example of Very Regular Service	173
5-12	3/11/02 Red Line Example of Mildly Degraded Service	174
5-13	Average Passenger Wait Times (minutes) in Evening.	175
A-1	SQL Code Used in Initial Database Access	201
A-2	Garage Check Oracle SQL Code Fragment	202
A-3	Position History Table and Index Creation	203
A-4	Code Fragment for Position Transfer to CTA Database	203
A-5	Usage of Database Position History Data	205

List of Tables

4.1	CTA Application Shared Information Needs Summary	83
4.2	CTA Application Choice Matrix	98
4.3	AVL-Available CTA Application Choice Matrix	99
5.1	MBTA Application Shared Information Needs Summary	147
5.2	MBTA Application Choice Matrix	161
5.3	Red line scheduled headway and ideal average wait times	166
5.4	AM Peak Periods With Avg Waiting Exceeding 110% of Minimum .	172
5.5	Evening Periods With Avg Waiting Exceeding 110% of Minimum . .	174

Chapter 1

Introduction

In this thesis I investigate the potential for cost-effectively utilizing the systems used in daily transit operations as data sources to be used agency-wide to improve line performance, agency efficiency and user information. Many agencies specify and acquire stand-alone devices to perform prescribed tasks, such as Automated Vehicle Location systems (bus), Automated Transit Supervision systems (rail), and Stop Annunciation packages, without sufficient consideration of the possible uses for the data these systems may provide for operations management and analysis, performance auditing, scheduling and strategic planning.

1.1 Objectives

In this thesis I use two case studies of major transit agencies to show that there is potential for significant subsidiary benefits for an integrated approach to the use of automatically collected data. This opportunistic use of existing agency data sources can provide an answer to a number of questions of practical importance to the agency, including:

- How can an agency more effectively monitor bus operator performance?
- What factors contribute to poor service reliability in typical operations?

- How can an agency predict developing service degradation before it becomes unacceptable?
- What techniques are effective at maintaining reliable service given this new understanding of line performance?
- How can an agency audit rail line reliability?
- What data collection technology and practices are of the greatest use to an agency?
- Finally, what are the detection cost and performance trade-offs for different levels of data availability and quality?

1.2 Motivation

There is significant pressure on transit agencies to reduce operational expenses. At the same time there is also significant pressure to provide higher quality service to the traveling public. Often these goals are at odds with one another. However by creatively using available operations data for performance auditing, scheduling and strategic planning the agency has the potential both to reduce operational expenses and to improve the quality of service.

Agencies traditionally have used in-house staff or contractual arrangements to collect key data manually. One example of this is the use of fixed location and mobile supervisors to monitor terminal and en route schedule deviation. Another example is an agency contracting out segment level run time checks for fine tuning schedules. Further information on this is provided in Section 2.1. Lastly, the ride and point-checks to monitor passenger load along a route are also typical recurring data collection activities. Manually collected data such as surveys, ride checks and point checks provide the agency with much needed data about their service and their customers. Unfortunately due to the costs associated with these collection efforts, many transit agencies feel constrained to collecting data on each of its routes

annually or less frequently.[5] This restricts what agencies can do to maintain high quality service, as significant changes in route performance may go undetected for some time. The quantity of data collected on each route may be suboptimal in a variety of ways. In some cases data are only available for trips on a single day. Even carefully designed manual collection programs rarely are sufficient to determine with confidence the variance of the characteristics of interest. Optimal running time and recovery time determination require not only the mean but also the variance of the running time to ensure a desired fraction of vehicles arrive on schedule and are available to start the return trip on schedule. Further, manual checkers are easily observed by the operator, and the operator may behave differently when the route is being checked. More extensive data is sometimes available manually from operator logs or supervisor logs, however the reliability of this data is suspect as it is being collected by someone with an inherent conflict of interest.

Technologies that are increasingly commonly in use can provide extensive data about agency service and its' customers. Data may be extracted from Automated Fare Collection [AFC], Automated Passenger Counters [APC], Automatic Vehicle Identification [AVI], Automated Vehicle Location [AVL], bus Stop Annunciation packages, (rail) Automated Transit Supervision and Control [ATS/ATC] and Computer Aided Dispatch [CAD] systems. Section 2.3 provides more detailed information on each type of system. Generally these systems have significant initial cost, but recurrent costs are typically limited to maintenance and data storage. This offers the potential for less expensive data collection while providing the large sample sizes needed for better characterization of operations than traditional manual collection methods. In-house staff assigned to collect data can be reassigned to focus on analysis of the data for such functions as service and operations planning. Service and operations planning is only one of a number of potential users of high quality data; there are several ways the use of automatically collected data can provide opportunities to improve the quality of service while reducing expenses. When the results of systems to audit line reliability, monitor bus performance, and understand line characteristics over time are made available to the operators, supervisors and controllers there is the potential to

increase headway regularity. By the principal of random incidence, when service is irregular more passengers arrive during long gaps in service than during very short headways. An increase in headway regularity will reduce the wait time and decrease the crowding experienced by the average passenger.

1.3 Approach

My approach starts with the development of a framework for the appraisal and fulfillment of agency needs using data available from installed agency data collection systems, as well as from data collection systems that may be installed in the near future. This framework is applied to the Chicago Transit Authority [CTA] bus department and the Massachusetts Bay Transportation Authority [MBTA] rail department in two case studies.

Framework application is a two step process, first the agency appraisal and second the application design and implementation. In the first step I investigate the potential for data collection systems that are currently, or may soon be, installed to help fulfill each of a number of agency needs, such as service quality auditing, passenger information, marketing. I select the most promising applications, and in the second step discuss critical decisions required in the design and implementation of the applications.

The three specific tools developed in the context of this framework use data derived from existing agency data collection systems to provide passenger information, improve service management and operations planning, and also provide a mechanism for performance auditing. Each case study concludes with a discussion of application effectiveness, and the lessons to be taken from the applications' varying degrees of effectiveness in improving agency and user decisions.

1.4 Thesis Organization

The second chapter lays the foundation for this research, reviewing prior work in this field. It begins with background information on agency functions as well as the types of information needed to perform those functions. This includes both specific data needs in transit agencies, as well as the needs of the traveling public. It continues with information on what data may be available from the data collection systems installed in many agencies, as well as the fundamentals of the technology behind these systems inasmuch as the technology impacts their usefulness. Next comes information on the processing of this data. Last it addresses traditional barriers to its use agency-wide.

The third chapter proposes the approach used to show that there is indeed potential for significant subsidiary benefits of an integrated approach to the use of automatically collected data. I present a framework for the appraisal and fulfillment of agency needs using data available from current and potential agency data collection systems. This framework is split into two sections. First is the agency appraisal: identification and selection of the most promising applications for a given agency. Second the design and implementation of these applications: what can be drawn from these proof-of-concept implementations.

The fourth chapter presents the case study application to the CTA bus department. The agency appraisal identifies two promising applications to be implemented and then the framework is applied to each application separately, addressing the issues of design, implementation and evaluation together. The first application addresses the use of operations data as a source for both real time passenger information and service management. This is drawn from experience at the CTA creating a system to provide web-accessible vehicle position reports using data from an on-vehicle emergency communications system installed on CTA buses. The second application involves the use of archived data for operations planning. This is a system that analyzes vehicle position over a long period of time to provide insight into line characteristics for use by agency operations staff to improve line performance.

The fifth chapter presents the case study application to the MBTA rail depart-

ment. The agency appraisal step identifies one promising application for implementation. This provides lessons on the use of operations data for service performance monitoring and auditing. The application was created in cooperation with the MBTA to combine train position information and historical demand data to provide management with daily line performance summaries from the passengers' perspective. This application provides a mechanism for supervisors to monitor system performance and gauge the effectiveness of interventions designed to restore service after a disruption.

The sixth and final chapter presents my conclusions and identifies future work in this field.

Chapter 2

The State of Practice

This chapter begins with a review of a range of functions performed within transit agencies. For each function this section describes the types and characteristics of the information needed, as well as the manual techniques that have historically been used to fulfill these information needs. The information needs of the traveling public are addressed next. This leads into an introduction to the range of automated data collection systems available and the roles in which they are typically used in agencies. The next section explores the technologies involved in automated data collection systems typically found in transit agencies, as knowledge of technology limitations is crucial when considering the use of this data in novel agency applications. Synthesizing information from the preceding sections, the next section explores the ways in which automated data collection systems may fulfill agency data needs. It continues by exploring integrated uses of this data, including those beyond the original specifications and intent of the agency data collection systems. This section includes a discussion of current state of the art in data processing tools with emphasis on the distinctions between this work and what has come before. Lastly, it explores the practical barriers to use of automatically collected transit data in the past and in the present.

2.1 Agency Needs

This section reviews the importance and describes agency benefits from good execution of each of a range of agency functions. This section also describes the types of information needed for the agency to carry out each function, as well as reviews the manual data collection techniques often used to fulfill these needs.

These agency functions are listed in rough order of increasing allowable delay between the collection of data and the availability of that data. The security, service management, passenger information and performance monitoring functions require very recently collected information to be effective. These are users of “real-time” data sources, requiring either direct observation in the field or information to be transmitted over distance with communications either by radio or by wire. The remaining applications, maintenance, marketing, operations planning, performance auditing, scheduling, route planning, and budgeting are each insensitive to substantial delays between data collection and effective use. These tasks are “post-processing” tasks, which can take advantage of more detailed information that is gathered over a period of time, with the information recorded and stored in a device for later download and use.

2.1.1 Security / Safety

Operator and passenger safety is of great concern to any agency. This need is strong enough to justify the purchase of dedicated on-vehicle systems in some agencies. In general, security consists of being able to make a call for help and communicate a vehicle’s position to someone who will dispatch assistance. A flag indicating distress, and the vehicle position are the only items needed for this task. This is traditionally provided by the operator making a verbal distress call including position over a radio link to a dispatcher or control center.

2.1.2 Service Management

Service management, or the restoration of normal service after some sort of disruption, is a very important agency task. Service disruptions occur in both major and minor forms. Major disruptions can be caused by, for example, drivers not showing up for their shift, vehicles breaking down en route, portions of a route or line made impassable due to weather or emergencies. Minor disruptions can be caused by, for example, increased dwell times due to load fluctuations, vehicle bunching [where multiple vehicles catch up with one another and travel together], or an operator showing up late for a shift. Without supervisory intervention, each of these disruptions will cause unreliable service, leading to increased trip times as well as crowding experienced by the user. If an agency with severe reliability problems, such as bus bunching, has set an acceptable average level of crowding experienced by the user, or set a likelihood of being denied boarding due to overcrowding, then by improving service management this agency can provide the same level of service with fewer vehicles, at reduced cost, or use those vehicles and drivers to provide service to new markets.

There are a number of interventions that a supervisor may consider. This includes short turning, which is turning a bus, or reversing the direction of a train before arriving at the terminal. In bus operations supervisors have a fair amount of flexibility in choosing where to short turn, while in rail short turning options are constrained by the available crossovers – connections between tracks. This will be considered when there is a gap in service approaching a section of a line or route with a high boarding demand, and there is a [usually lightly loaded] vehicle travelling in the opposite direction in the vicinity of that high demand point. This can be a very effective method of service restoration. Expressing, where a [usually heavily loaded] vehicle does not make its usual stops for some portion of its route can also be used. This has the potential to shrink a gap in service by allowing a vehicle to catch up with its schedule, but at the expense of passengers travelling to or from the skipped stations. Another option is holding a vehicle, which can be an effective means to re-space vehicles and provide more uniform headways, improving service for users down

the line. This also has the potential to greatly reduce the impacts of disruptions. [2] In some situations, primarily in bus operations due to the ease of changing routes, a vehicle may be interlined to fill a gap in service. This can be especially effective in replacing service on low frequency routes where a missed run can lead to a very long service gap for users.

Data required for service management include the vehicle position as it relates to progress along a route, for all vehicles on a route of interest. If the headways on a route are long enough that a substantial fraction of users time their arrival to the schedule, then schedule adherence – the comparison of position-on-route to the schedule – is also useful for each vehicle. Information in real-time on actual vehicle loads would be ideal for selecting likely vehicles for interventions. Historical demand and load profile information – summaries of boarding and alighting demand – are often more readily available and can be useful for service management, eg. deciding where to short turn a bus based on inconveniencing the fewest passengers.

Up-to-date position and vehicle load information historically comes from supervisors on the street or in stations and towers, typically in communication with a central dispatcher or controller. Historical demand information is typically collected via ride-checks – employees who are paid to ride the length of the route noting times and passenger movements.

2.1.3 Passenger Information

Passenger information is important to the agency to increase demand and provide a more satisfying experience for the user. Demand increases are typically from users with a choice of modes, who will take transit if given sufficient information. The data needs of the users are addressed in Section 2.2.

The minimum passenger information is an indication that transit vehicles will stop at a particular location – a bus stop sign for example. More useful signage will have the route name, and perhaps the route structure and information about the span of service. Informing users of typical headways for a given time of day can be useful for trip planning. Depending on typical service reliability, the route schedule

can also be helpful. Real-time information for passengers, such as are used in service management [vehicle position and load for each vehicle on a route], provides an even higher level of service. The information provided to the user needs to be consistently accurate in order to earn the users trust.

Static data for passenger information is typically taken from the route structure and the schedule. Manual real-time passenger information is sometimes relayed over loudspeakers in rail stations, and sometimes the crew on a crowded train may provide manual real-time passenger information informing passengers how long it will be until the following train arrives.

2.1.4 Performance Monitoring

Performance monitoring is typically done intuitively in the course of service management. Automated tools to support this function can help improve the performance of new controllers and dispatchers, by acting as an automated reminder of the quality of service experienced by passengers. Such a system can also have a further degree of automation and be part of a computer aided dispatch system, where the system suggests vehicle departure times as a further aid to dispatchers. Data required for performance monitoring is the same as required for service management.

2.1.5 Maintenance

Maintenance has substantial capital costs to the agency. Also, vehicle maintenance influences service reliability, which, as described in Section 2.1.2 means Maintenance has an effect on route performance. Vehicle performance indicators, such as engine temperature, odometer reading and any trouble codes are all useful pieces of information for monitoring vehicle health. The health of other on board systems can be monitored by looking for anomalies in any given vehicle's data outputs.

Some of this functionality has typically been provided on certain vehicles by reading a stored record of trouble in the engine computer when a check engine light is illuminated, however this does not provide information about ancillary systems health.

Potentially new data collection systems, such as those that continuously record on-vehicle or wayside noise can identify maintenance needs before they degrade to a point where an operator files a complaint. Such a system could detect flats on rail wheels, for example, on a particular train, within minutes of the hard braking incident that caused the problem. Affected vehicles could automatically be enqueued for repair, and taken out of service as soon as practical. This has the potential to reduce further damage to vehicles and tracks caused by a malfunction.

2.1.6 Marketing

Marketing is a form of non-real-time passenger information, but it typically has a much wider intended audience, and can potentially attract a large number of users. Data used for marketing may include on-time performance figures, if especially improved, overall journey time, if especially competitive, or more typically static information such as increased hours of service or the opening of a new route or rail line. Variable message signs on the highway could provide transit public service announcements, which utilize transit statistics from the previous weekday, or the previous hour. For example, a highway could have a sign presenting users with actual times to a central business district from a nearby park and ride station. This could be especially effective if only presented in periods of highway congestion.

2.1.7 Operations Planning

Operations planning encompasses a number of agency tasks related to how service is provided. One such task is scheduling, which will be addressed separately in a following section. Planning for special events is another key part of operations planning. This includes the consideration of service management strategies to be implemented should service start to degrade. This is an important agency task which has the potential to improve responses to certain undesirable situations which may occur relatively frequently. For more information on this topic see [25].

The data required for operations planning include the same data items as for ser-

vice management, with the addition of other possibly explanatory information. For example, recording operator badge number, as well as the vehicle type, to facilitate identification of root causes of unreliable service. More important for special event planning, the maximum throughput of certain elements of the system can be measured. An example of this is the maximum number of trains that can run past a given point on a line in a half hour period, and the sensitivity of this figure to demand.

2.1.8 Performance Auditing

Performance auditing is post-processing of measures of performance to provide an indication under what circumstances operations planning and service management do well, and under what circumstances they do an inadequate job of maintaining or restoring reliable operations. This function can be used to highlight particularly effective service management practices as well as discourage the use of less effective interventions.

Using automatically collected data, a great deal more can be done to measure service performance. These measurements may be used to identify the specific situations in which operations planning can be improved. There are a number of route-level figures that can be used to compare line performance over a series of days. A simple figure is the percentage of vehicle-time spent at less than half the scheduled headway behind its leader. This figure can give a sense of how often vehicles are catching up with their leaders and remaining bunched. An advantage of this figure is that it is possible to calculate from data that may be incomplete. If entire vehicle runs are missing from the data set, this figure will still be useful. For vehicles which are successfully tracked and are excessively close to one another, you can predict that a following vehicle will have a longer than usual headway – which is what the customer experiences as a poor level of service. This may tend to underestimate the passenger impacts in a system where runs are often held in, due to mechanical problems or absenteeism. Using percentage of vehicles which are bunched may tend to overestimate the impact if demand is low enough that following vehicles can stay on schedule.

In a system with more reliable data, this uncertainty may be removed by using figures more directly related to passenger waiting. The simplest of this class of figure is the percentage of vehicle time spent at greater than two scheduled headways from its leader. If demand patterns are known, an even more customer-focused figure, such as expected average waiting time can be determined, and compared from one day to the next. While average passenger waiting time is good for route-to-route comparisons, and can show variation in day-to-day comparison, the percentage of passengers waiting longer than a threshold value is potentially more valuable for deciding whether an intervention worked well. If a small number of passengers are treated especially poorly, while a larger number receive high quality service, this may not be in the best interest of the agency, as it may permanently lose those poorly treated customers to an alternate mode. If this happens repeatedly, it can erode the user base. A system with extremely conservative policies on scheduling and holding to schedule may benefit from a more all-encompassing figure of average total-trip time. This should be weighted according to local measures of passenger impacts of out-of-vehicle waiting time. This would take into consideration both the reduced delay in boarding from increased reliability, as well as the increased trip time caused by operating the vehicle slower than its capability. It is up to the agency to decide on a focus, and which measure or measures are most appropriate.

Performance auditing uses similar information to performance monitoring. In addition, segment-level running time, which is a further derivative of position-on-route information can be useful for identifying vehicles traveling unusually fast or slow on each segment. Vehicle position data at finer increments can be used for more in-depth performance auditing, analysis of dwell time, traffic impacts, operator effects and so on. Performance auditing often uses load at a peak point, which is a subset of the previously mentioned load profile. Given vehicle position information to infer headway, the historical load profile can be used to provide loading estimates for vehicles. Ideally a performance auditing system should also have a record of all the incidents and interventions on the line, to separate these from delays caused by dwell-time or traffic. On-time performance of vehicles at the terminal as well as en-route

at time points is also useful for monitoring driver performance.

Information for performance auditing has historically come from point-checks – an employee at a fixed point tabulating passenger loads and arrival time. Terminal and street supervisors have also traditionally had a role in auditing the on-time performance of the vehicles on a route.

2.1.9 Scheduling

Scheduling is important for the reliability of service on a route as it determines the amount of layover time a vehicle is allowed before it starts a return trip. If the layover time is insufficient, service quality will likely continue to degrade as the vehicle becomes more heavily loaded by picking up passengers from a longer gap from the preceding vehicle. Scheduling too much recovery time can be costly and can also delay passengers on approach to the terminal if there is insufficient terminal capacity to accommodate all the lay-over vehicles.

Scheduling uses much of the same information as performance auditing, including segment-level running time as a further derivative of position-on-route information for setting time points along the route. Scheduled headways are often set to achieve a certain number of passengers at the maximum load point over a period of time. The peak load is needed on a number of vehicles to assure compliance with this policy. More aggregate measures of passenger demand, such as total boardings for a route and revenue for a route are also used for headway determination. Extensive data sets are invaluable in setting recovery time to assure a particular percentage of vehicles are available to start a trip on-time.

The information on segment level running time and load profile used to determine the peak load point are typically collected manually through ride-checks. In some agencies operators keep track of total boardings, and revenue information is determined when the vault is emptied at the end of the day, or may be checked more frequently for a special data collection. Recovery time has traditionally been set based on rules of thumb as well as operator feedback.

2.1.10 Route Planning

The path a bus route or rail line takes through local neighborhoods can have a large influence on the demand generated. Reducing route duplication can decrease costs with a minimum impact on service quality and coverage.

Aggregate measures of passenger demand, such as total boardings and revenue for a route are used in route and system planning along with more detailed information about passengers, such as information on revenue by fare type per route, as well as transfer rates between routes. Transfer rates can be used to estimate the total number of individuals served – the system-wide linked trips figure. This is a better proxy for how effectively the system serves the public. The expected demand for service after proposed changes in bus routes or the addition of new rail alignments can be predicted using the underlying passenger origin and destination demand information, as well as demographic data.

Transfers can be manually tabulated by collecting transfer tickets at the end of a trip. Passenger origin and destination demand information is collected infrequently, but when collected it is often by passenger survey.

2.1.11 Budgeting

Budgeting, and especially government funding, is important to all agencies. Funding is, in the US, linked to data collection.

Counting each new fare or transfer, demand is aggregated system-wide to provide a count of unlinked passenger trips. Average trip length is also a figure of merit for transit systems which can be determined for a route from the load profile. This figure combined with total boardings for a route is summed to provide system-wide passenger-miles. In the U.S., system-wide figures such as passenger-miles and unlinked trips are reported to the Federal Government for inclusion in the National Transit Database as a requirement for federal funding.

2.2 User Data Needs

Three distinct categories of prospective passenger needs are pre-trip, at-stop or in-station, and on-vehicle needs. Pre-trip, users are concerned with choosing a route and the time of departure needed to minimize waiting or to arrive at the destination at a given time. In a rail station or at a bus stop, the user needs to know how long he or she will have to wait, as well as route structure and other information to aid in dynamic decision making. On board a vehicle, the user needs to know when to alight, and would also like to know trip time remaining.

While the transit agency itself needs information for “passenger information” purposes, the underlying passenger needs which ostensibly drive the agency’s procurement are often not considered explicitly. At a minimum, users need to be informed of the route structure. Better still, information on span and frequency of service should be provided. A further improvement is scheduled arrival times for specific trips, at least for long-headway service. All of the above needs are typically met using printed system and route maps. Some agencies, such as the CTA, provide a phone number on all bus stops that will connect the user to customer service agents to provide all of the above information. As users and agencies become more computer savvy, World Wide Web based schedule information distribution and trip planning is becoming more popular. [32]

The next level is real-time information about actual system conditions. In a number of rail systems this is provided manually in stations and on vehicles via public address systems. However this is typically only under exceptional circumstances such as an interruption in service. Dynamic information about normal operations is often available on the vehicle by driver announcements at stops. This is more reliable – though less critical – on rail systems due to the reduced number of stops and clearer station signs. In-station, on-street or pre-trip real-time information about normal operations, such as the time of next vehicle arrival, is almost exclusively handled by automated systems. More detailed information involving both vehicle and passenger demand in real time would ideally be available wayside, such as time until the next

bus with available seats.

2.3 Automated Data Collection Systems

A wide range of automated data collection systems are now becoming more generally available. Many are used in a stand-alone configuration, and this section will describe the typical uses for which these systems are designed and deployed.

2.3.1 Security / Emergency Location

Agencies require a mechanism for the operator to quickly summon help in cases where the operator or a passenger's safety is threatened. Automated position reporting systems with a "panic button" silent alarm feature are augmenting or taking the place of the traditional voice distress call. These vehicle security systems' primary responsibility is to automatically summon help to the vehicle without arousing suspicion. Operators are often required to enter some identifying information into the system at the beginning of the route, such as badge number and some can provide text messages to the operator. These systems typically use GPS-derived position information as they may not be at a location where odometer readings have meaning or signposts are available.

They require a radio data link, and could potentially benefit from the lower communications delay of a request-for-channel type system (see Section 2.4.2), as information is exchanged infrequently, and there is a premium on minimizing delay. On the other hand, the CTA bus Emergency Location system is an example of a system using polled data channels. In some systems the emergency location hardware is part of an in-vehicle surveillance package which automatically records the last several hours hidden camera footage, and some have the ability to transmit a still image back to a control center or security office.

2.3.2 Automated Vehicle Location

AVL systems combine one or more location technologies (see Section 2.4.1) with a (polled) radio data (see Section 2.4.2) link to provide vehicle location information on a regular basis, along with some identifying information, and some vehicle status information. The identifying information is often manually entered into the AVL device, including one or more of the bus number, the run number, the route number and the operator badge number. Vehicle status information may include the door open/closed state, and perhaps engine temperature, odometer reading or other vehicle condition information.

These systems do not typically store or process data on-board. They usually simply return location as coordinates or the identity of the last signpost passed, without attempting to use route structure information to map that to a position along a route or a deviation from schedule. Data are typically used in the control center for service management – for example to identify candidate vehicles which may be redirected to restore service after a disruption. In some systems the data are stored for a limited period of time to help with incident investigation. Also in some systems, position is processed on board along with schedule information to provide an indication to the operator about current performance compared to the schedule.

2.3.3 Automatic Vehicle Identification

AVI based location uses wayside, signpost type systems where vehicles have a tag or token from which an identifying number is remotely read at some distance from the vehicle and at speed. Such detectors can be used to automatically set routings in branching systems. Data from such systems can be used for control center display of vehicle position but can be subject to missed detections. If this type of data is to be used effectively for service management, a substantial number of detectors must be placed along a route to provide adequate coverage for centralized line management, at least at key points along the route such as the ends and at a distance out on the branches. Trains on the MBTA's Green Line light rail system are routed using

thirty-one AVI detectors along the quadruple branching route. While the AVI detections are displayed in the MBTA Operations Control Center, much of the actual line management is done in the field. Recorded data are used for incident analysis.

2.3.4 Automated Train Supervision and Control

ATS and ATC systems are both exclusively rail technologies where the vehicles are tracked primarily using segment level detection by power controllers. The presence of a train on the track completes a circuit and this signal is digitized and sent to a control center. ATS systems include a computerized display of a range of line information, including the track occupancy, signal, switch and start light information. [33] ATC systems have the added feature of having some direct control over the vehicle, at a minimum enforcing restrictions on how close one vehicle may be to another, and at what speed. More advanced systems may include schedule information and have control software to attempt to keep vehicles on schedule automatically. These systems may be augmented by AVI detectors and or dead-reckoning wheel rotation sensors, to provide position at a sub-track-segment level. The MBTA in Boston and Tren Urbano in San Juan use ATC systems albeit of different capabilities. As in other vehicle location systems, data is recorded for incident analysis.

2.3.5 Trip Time Analyzers

Trip Time Analyzers are designed exclusively for post-processing and analysis. They log detailed information about vehicle position, performance and progress along a route to an on-board computer from which data is downloaded nightly for analysis. These systems are not constrained by the radio bandwidth restrictions of real-time communications. Rich data are recorded, including each time a vehicle stops, whether the doors opened or not, for how long the vehicle was stopped, periods of time where vehicle speed is below a threshold, maximum speed since last stop, and the time and location of passing signposts or “virtual signposts” – positions of interest along a route for which the computer will detect and trigger an event. Nearly any other location,

speed, or direction-based measure that is desired can also be recorded. Some of these systems have a processing component with a copy of the route schedule. This is used to automatically compute schedule deviation which is presented to the operator.[26] These systems are designed to provide data for analysis after the fact, and data are carefully archived for that purpose, including as much supporting information as possible, including operator, vehicle, run and route identifying information.

2.3.6 Automated Passenger Counters

APC are Trip Time Analysis systems with extra sensors for counting passenger boardings and alightings. A variety of sensors are in use, including break-beam sensors, pressure sensors, infrared sensors and video sensors. Break-beam sensors track the movement of passengers by the sequential interruption of two or more beams of light, for determination of direction of travel. Pressure-sensitive mats can be simple or complex, with simple systems having a minimal number of pressure sensors, two or three detecting footfalls on entry stairs. As low floor buses become more prevalent, more sophisticated sensors are being used which utilize an array of small pressure-sensitive “pixels” that are used to track individual feet, and can distinguish and track multiple sets of footprints. Overhead infrared sensors are only useful in climates where body temperature is greater than the ambient temperature. Overhead video cameras are not yet in widespread use but, along with the sophisticated pressure sensors, have the potential to count on rail cars and buses with high throughput entry and egress.

Unlike the previously listed systems, APCs are often only purchased for a fraction [typically 10% to 25%] of the fleet, and they are rotated through different routes on the system. A partial fleet equipped with APCs is sufficient for planning and productivity reporting, leading many agencies to save on the purchase and maintenance costs associated with equipping the full fleet. They do have to deal with the added hassle of distributing the APC buses on a representative sample of all systemwide trips. Further, if a detailed study of a single route or set of routes is desired, they need to get blocks of particular vehicles on the same route, which can be quite difficult if there is a mix in the garage.

2.3.7 Stop Announcement Systems

US agencies are required by the ADA to make bus stop announcements. Such announcements are also useful for infrequent and first-time travelers on a given route. Automated stop annunciation [SA] systems have been developed which provide this capability without operator involvement. These systems are effectively Trip Time Analyzers, without the data logging and with an added passenger information display. Just as in Trip Time Analyzers, the SA system requires route structure information to know what stop the bus is approaching, but it is simpler since it does not refer to scheduled times, nor does it calculate time deviation for the operator. The SA system requires accurate location information, so unaugmented GPS is insufficient (see Section 2.4.1). While these systems do not typically collect data on their own, agencies may have the opportunity to install SA systems with a data collection capability as a low-cost added feature.

2.3.8 Electronic Fareboxes

Electronic Fareboxes, also known as Automatic Fare Collection [AFC] systems are, for better or worse, the only data collection system that is typically ignorant of the vehicle location. AFCs typically record use by fare category by time period. The more advanced AFC systems are transactional – each fare collected is time-stamped for future analysis. These systems are in wide use for providing revenue and total boardings on a route-by-route basis.

2.3.9 System Integration

As more digital systems are installed on vehicles, it becomes more important to require standard interfaces when initially specifying systems for purchase. Some modern buses have Electronic Destination Signs that are independent of the AFC and the AVL system, each of which requires a separate operator login. As more logins are required, the chances of operator error increase, and achieving the benefits possible via coordination of systems' data become increasingly difficult. A number of on-vehicle

data communications standards are available and will be discussed in Section 2.5.

2.4 Technology

Location and long distance communication technologies are used in many of the automated data collection systems listed above. The constraints of the technology used on vehicles often influences what the data may reliably be used to accomplish.

2.4.1 Location Technology

Most automated data systems, including Silent Alarm security systems, Automated Vehicle Location [AVL] systems, Automated Passenger Counters [APC], Stop Announcement [SA], Automatic Vehicle Identification [AVI] systems, rail Automated Transit Supervision and Control [ATS/ATC] systems, and Trip Time Analyzers have an automated location component. The location of the vehicle is determined using a Global Positioning System [GPS] receiver, a dead-reckoning type system, some sort of infrastructure [signposts] at key points along a route, or a combination of multiple technologies. The class of in-vehicle automated data systems in use that typically does not have a location component are Automated Fare Collection [AFC] systems also known as Electronic Fareboxes.

GPS

The GPS-based approach involves a receiver that can interpret signals from four (or more) satellites to provide a three dimensional position estimate that can locate the device anywhere on the planet. Standard GPS accuracy in the horizontal plane is 15m or better 95% of the time.[11] Several mechanisms for improving on this accuracy have been developed, including differential GPS [dGPS] which yields 5m accuracy 95% of the time. This takes advantage of the fact that GPS errors on nearby devices will be strongly correlated, and error can be reduced in a mobile unit by broadcasting corrections from a fixed unit at a known location. The latest advance is the Wide Area

Augmentation System [WAAS], a system of geostationary satellites which promises 3m or better positioning 95% of the time.

However, those figures come with a number of caveats. First and foremost the accuracy numbers provided are for devices with a clear view of the sky. The signal will be lost when the vehicle is in a tunnel or even a terminal building with a metal or concrete overhead structure that blocks the view of the sky. In an urban environment with tall buildings, satellites will go in and out of view, which can lead to sequences of estimated positions which imply the vehicle experienced impossibly great acceleration as different satellites with different errors are used. Even more problematic for transit agencies using GPS is the problem of multipath. GPS signals reflect off large flat surfaces, such as buildings, and may be detected as having traveled a longer path than they actually have. While a GPS receiver can detect that it is not seeing sufficient satellites for navigation and “coast”, in the case of multipath, it is deceived by the reflected signal and does not know there is a problem; this may lead to readings hundreds of meters off from true, and in some cases with plausible direction and velocity components. While in an aviation system usability and reliability can be improved using a Local Area Augmentation System [LAAS] – effectively a local, fixed position transmitter that appears to the vehicle to be a GPS satellite – this will not help in the urban environment. See 2.7.6 for further issues with GPS use internationally.

Dead-Reckoning

Dead reckoning [DR] is a general term for a class of navigation where an initial position is entered into a system, and from then on it keeps track of its location using only sensors that observe the state of the vehicle. Given sufficient hardware this can be an extremely effective method of position determination; submarine navigation systems can achieve drift of a few meters per hour using just velocity sensors and inertial sensors.[20]

The least expensive DR system uses information already on-board a vehicle, such as turns of a non-driven wheel [odometer] and, potentially, angle of the steered wheels.

Such systems can be effective if the route is known and the device is calibrated frequently. Of course, drift can be a problem, especially when the pavement is uneven. Such a system does not generally accommodate detours gracefully although if steering wheel position is tracked, detours can be identified, if not accommodated. One clear benefit is that such a system does not need a view of the sky and hence can be used in tunnels, stations and urban environments. Subway vehicles often use signpost-type infrastructure-dependent positioning, but some, including the new Siemens trains in the Tren Urbano heavy rail system use wheel revolution detectors for more precise positioning.

Inertial Navigation Systems [INS] have been used in research prototype vehicles[39] and as technology prices drop may not be out of reach of transit agency budgets. INS uses accelerometers to sense change in velocity, which is integrated to estimate velocity, which, in turn, is integrated to estimate position. More sophisticated DR systems may combine odometer and inertial information to provide a superior position estimate.

Signpost

Signpost based AVL systems are characterized by the installation of hardware along the route that “places” the vehicle at key points in its trip. The electronics can be on the vehicle with passive route markers, as is typical in buses, or in the wayside infrastructure with passive vehicle markers, as is often the case with trains. Signpost data have the advantage of being available at exactly the points of primary interest, without requiring any interpolation. It also works in situations where GPS-based ranging systems are unavailable, similar to dead-reckoning. There is some loss of route flexibility in a system that requires wayside hardware, but it recovers from short detours more gracefully. Rail ATS and ATC systems provide a form of signpost-type information; hardware in track-side bungalows identifies vehicle position by electrically sensing train presence on a track segment.

Combinations of methods

Combinations of GPS, dead-reckoning, map-matching and signpost technologies can be extremely effective, as the deficiencies in one technology are often strengths in alternate technologies. For example, odometer based route positioning tends to become less accurate over time as error accumulates, however even infrequent augmentation with a signpost-type system can correct the error in the position estimate. GPS systems augmented with odometer readings and direction sensors are also extremely effective at maintaining a good estimate of vehicle position, since the expected trajectory [the route] is known in advance.

2.4.2 Communications Technology

The main division in communications technology in use is between the fixed rail infrastructure and the mobile bus platform. Rail systems primarily use wire data links to a control center for real time communications. Such systems vary in their communications mechanism, but modern systems such as the MBTA's are wired using long-distance serial data links between a control center communications processor and bungalows in the field with a digital electronic track power control system. Older systems, such as much of the CTA, do not digitize the rail power control information [track on/off data] in the bungalows, but instead run a line to the control center where the single bit of information [track on/off] is digitized. The latter method is an inefficient use of communications capacity, and is prone to error.

Bus systems require using an over-the-air radio data link to send and receive information. When sending data over the air, agencies have to choose between licensing dedicated radio-frequency [RF] spectrum from the FCC, and setting up antennae over the service region, or utilizing one of the public cellular telephone data networks. Use of the public network is attractive due to low start-up costs and relatively high bandwidth available, but the per vehicle operating costs are high. Mitigating the attractiveness of low start-up costs, many agencies also already have a substantial dedicated radio infrastructure they may migrate to data communications at minimal

cost. For these reasons and others, agencies typically utilize dedicated licensed channels. In such systems a number of vehicles must share each data channel since these channels are a scarce resource. An agency may be able to distribute its vehicles over a number of channels for added bandwidth. The mechanism for sharing the data channel is typically a procedure where the vehicles are polled on a regular basis, which is a repeated query for any new messages they need to deliver to the control center. The CTA uses such a system with a poll cycle between 60 and 120 seconds for vehicles on different channels. An alternative arrangement is a request-for-channel broadcast from the vehicle: digitally emulating a voice hail. This has the advantage of reduced delay between the vehicle operator request and the control center response, and is good for systems where vehicles do not usually have a message to report. However this does not scale well.

2.5 Interface Standardization

Standardization in the hardware and software interfaces of systems used on vehicle is becoming more common as agencies recognize the difficulties in integrating systems from multiple vendors, as discussed in Section 2.3.9. Nevertheless there is substantial pressure towards proprietary systems. For example the initial cost of proprietary systems may be artificially discounted by companies which expect to capitalize on a captive market for future enhancements.[10] Even when data interfaces are standardized, the transit industry faces the problem of having too many competing standards. This limits the extent to which agencies may mix hardware and software vendors.

There are a number of “standards” competing for industry and agency use on vehicle. In bus communications, a Vehicle Area Network [VAN] of information components using the SAE J1708/J1587 network is supported by some due to its proven robustness. The J1708 hardware standards and J1587 communications protocols were first developed for information exchange between drive-train components including the engine, transmission, and brakes on heavy trucks. It was adopted for bus information-level use in 1992, while drive-train components migrated to the faster

SAE J1922. Critics of J1708/J1587 VAN cite its slow communications speed as a reason to adopt a newer standard. Europeans are generally using the Controller Area Network [CAN] developed by Bosch. France is a proponent of a third protocol named WorldFIP. [tcrp43] Standards for rail communications have been accepted by the IEEE as standard P1473, which encompasses two distinct but compatible systems now in use: the LonWorks standard by Echelon, as well as the Train Communications Network [TCN] standard developed in Europe. In the bus domain, such on-bus communications networks are often referred to as a “smart bus” design.

Off vehicle, there are other standards defined in the National Transportation Communications for ITS Protocols [NTCIP] Guide. These standards are protocols to facilitate information exchange between agencies. They can be used within a transit agency for facilitating the combination of multiple data sources, as well as between transit agencies, for allowing software sharing, and between a transit agency and another agency, such as the highway department for signal coordination. US Government funded ITS projects are required to use this family of protocols, which should encourage its adoption. [15]

2.6 Processing of Automatically Collected Data

A large quantity of data is generated by any of the systems described in Section 2.3. These data products are in many cases used exclusively by software provided by the hardware vendor but often vendor software can also provide data sets for export which can be used in agency and external analysis. In other systems, a third party software company uses the hardware vendor data output specifications to provide enhanced tools.

2.6.1 Vendor Software

Vendors often provide, or sell, data processing tools for use with their hardware. Infodev is one of a number of hardware providers which produces data analysis software to complement its APCs. [18] Their BusStops software is typical in that it will pro-

vide a few simple figures about bus operations, including time-at-stop tables, route level load profiles and system wide figures. The software is strongly coupled to the hardware, implying that to continue using this software through a hardware upgrade may be difficult. Even within the field for which the hardware was designed – planning for APC, operations for ATS systems, etc – vendor-provided software often does not meet all the requirements of an agency. Agencies must set aside money to fund agency-specific requirements. Charges for changes beyond an initial contracted support period may be especially difficult for an agency to justify, leading to software with which the agency is unsatisfied.

One vendor that does attempt to provide an integrated software suite for use of agency data is INIT Innovations in Transportation, Inc, the US subsidiary of INIT GmbH, which sells vehicle, control center and passenger information hardware. [19] They also provide a range of software packages for a wide range of agency uses. Their package MOBILEstatistics is used for access to single events for incident investigations. This tool can be used for analysis of longer time periods for evaluation of operational procedures at a greater level of detail. A package called MOBILEreports is designed to help in system performance auditing, listing punctuality or route diversions. Their product STOPinfo is the user interface system component. They also promote their open interface, named open.P, which is the framework under which data are integrated. [8] One other company, PTV,[31] also uses this standard. While the open interface is important, having so much of the agency’s key data analysis done by software that is beyond their control is a risk to the agency. Clever Devices is another vendor which provides complete hardware and software solutions for agencies, with an Automatic Vehicle Monitoring product to compile important service performance characteristics. They have a tool called CleverTrack which is a management tool providing an interface to real-time vehicle data.

2.6.2 Specialized Analysis

Much data processing and analysis are done by the transit agency, or researcher, using data exported into general purpose tools such as spreadsheets and statistical analysis

packages. There are also instances of use of specialized, purpose-built, programs written by the agency or researcher to support operations and data analysis. There is little inter-agency sharing of house-built tools.

There is a significant body of work of researchers using archived data from vehicle position systems used in operations. In the field of service management, Eberlein[7] in 1995 using MBTA Green Line AVI data as key inputs to evaluate real time control strategies. Strathman, et al., in 1999 and 2001 used data from thousands of trips over 28 weekdays of service on a variety of bus routes in Portland Tri-Met's system to evaluate the effectiveness of headway-based control measures in the agency's Automated Bus Dispatching System.[36],[35]

Lam[21] provides a thorough overview of the technologies involved in providing real time passenger information as well as describing several transit agencies which deploy passenger information systems designed as an add-on to a control center where vehicle location data are already available in some form.

Another example of a system using automatically collected data designed to increase efficiency is the Passenger Waiting Time system implemented for the MBTA.[41] This system used primitive train location hardware to provide a real-time display of delays and schedule adherence information, as well as printed and faxed summary reports for system performance auditing. The resulting Passenger Waiting Time system was retired when the MBTA transitioned to AVI and ATS systems. This illustrates the importance of decoupling data collection and distribution from the processing and display.

A great deal of work has been done using archived vehicle position data in planning. The TCRP report "Data Analysis for Bus Planning and Monitoring" [10] does a thorough job of summarizing the current state of the practice, as well as showing that measurement error can be controlled and accounted for and, better still, that the large sample sizes available in automated systems for the first time provide agencies with the opportunity for statistically valid running time and schedule adherence analyses. It also points out that there is a lack of cooperation between agencies, and that data analysis expertise may not be available at many transit agencies. It discusses

difficulties in the use of AVL data for planning, but primarily focuses on the high quality data available from trip-time analyzers.

There is a significant body of work of researchers using archived vehicle position and passenger flow operations data. Horbury[17] used six days of archived signpost and dead-reckoning AVL data to estimate schedule deviation and patronage. Furth used AFC data from Los Angeles area bus routes[9] to provide a mechanism to estimate average trip length from extensive boardings data available from AFC systems. Rahbee[33] did substantial post processing of MBTA Red Line ATS data to make a series of recommendations intended to reduce passenger waiting time without additional system cost. Lee[22] used automatically collected data to explore operations at CTA terminals.

Because of the limited flexibility of proprietary packages for data analysis, a number of agencies use general purpose database and analysis software customized in house or with the help of outside contractors. OC Transpo in Ottawa has developed software in house for its analysis of APC data, including trip time analysis and report generation. The agency has used APC data for all its non-survey data collection needs since 1981 when its 8 traffic checkers were redeployed within the agency. [37] In December 2001 it started using Infodev's BusStops software to process data from Infodev hardware installed on light rail vehicles.

2.6.3 Third-Party Tools

According to Furth, [10] the use of third-party tools is increasing in agencies. This can reduce agency dependence on a single vendor by separating hardware and software procurement. It also has the advantage of reducing duplication between agencies. It may increase agency flexibility compared to vendor-produced systems, but agency-specific needs are hard to satisfy when use of third party software replaces, rather than complements, the hiring of technically savvy staff required for in-house software development and customization.

Some software packages that have originally been created for one purpose have been extended to accommodate automated sources of data. One such product is the

popular advanced crew and vehicle scheduling package Hastus by Giro. HASTUS-Rider, HASTUS-ATP and HASTUS-Vehicle are add-on packages to accommodate automated passenger count information and to use AVL data to fine-tune segment level run times that are used in scheduling.[12] Some software originally developed for analysis of manual surveys, point checks and ride checks can be modified for use with automatically collected data. APC data can be used by RideCheck Plus, a software package by RSM Systems,[34] which primarily produces tabular output, but also includes some graphical reports such as boardings and alightings by period. Integration with Geographic Information Systems datasets, such as stop inventory and census data allows stop-level and route-level data analysis. The APC vendor and software provider must be in close contact to assure data is available in the appropriate format for the software used.

The state of the art in off-line transit data processing is a tool called TriTAPT (Trip Time Analysis for Public Transport), developed at the Delft University of Technology.[26] This tool has been designed around the volumes of data provided by Trip Time Analyzers. TriTAPT has been primarily used with signpost based data, and is still in the process of being adapted to use the GPS based data that will open up the US market. This package can summarize per trip schedule deviations in an especially effective graphical interface, as well as provide means and standard deviations needed for schedule tuning. One major hurdle to adoption is related to the historical and even present day lack of standardization in agency data sources. The transformation of agency data to a format usable by this software is still up to the agency. Like most third party software it is closed source; agency customization must come from the vendor.

2.7 Barriers to Use

Historically there have been substantial barriers to the agency-wide use of automatically collected data. These include issues related to the availability of data in the appropriate form, the dissemination of information about presently available data,

and the confidence in the quality of data generated by these systems. This section addresses the extent to which these barriers persist.

2.7.1 Data Formatting and Archival

Even now in some agencies, operations data enters towers and control centers and is logged directly to strip-charts without being digitized. This is the case with track occupancy circuits in several of the Chicago Transit Authority's heavy rail lines, for example. In this environment, using operations data for planning and performance auditing has been impractical as the labor required to retrieve and reformat data from operations systems was better spent manually collecting data first hand in the field. This argument against use of operations data is becoming weaker and weaker as systems are replaced, with the data from the new systems always accessible from general purpose computers.

While systems data are increasingly available electronically, the stand-alone mentality remains. Electronic data access is often tailored for a specific purpose, such as incident auditing. Stored data is rarely matched with its route, or the associated vehicle schedule. Operator schedules and logs are also stored separately. When archived automated system data are not used for other reasons, then data ends up not being archived in a manner which facilitates future analysis.

2.7.2 Programmer Availability / Organizational Issues

Agencies in some cases lack the computer expertise in house to do the initial integration of various data sources. Furthermore the lack of in-house interest in data use has allowed continued procurement of isolated, stand-alone systems. Currently there are a number of hurdles remaining to implementation including the fixed costs associated with the initial setup of systems to bring operations data to a format that may be used agency wide. For example, agency personnel's time and in some cases the cost of new computer hardware are required.

In some cases there are organizational barriers to cooperation. The people who

work in the control center and are familiar with the data formats used may be physically separated from potential data users in the scheduling department.

2.7.3 Historical Lack of Standards

Efforts to standardize interfaces (see Section 2.5) will help simplify synthesis of disparate data sources and improve the portability of agency-created processing programs. However, there are still a variety of standards for on-vehicle communications. Furthermore there is still a great deal of on-vehicle hardware that was installed before these standards. Similarly, many existing agency data systems have data in a proprietary format, not conforming to the NTCIP family of standards.

2.7.4 Security Through Obscurity

A final objection to integration of disparate data sources is that some agencies protect sensitive internal data only by limited disclosure and documentation of the methods required to access this data. This is a risky strategy. While compartmentalization of sensitive items is good practice, it can and should be under the umbrella of an integrated, consistent system. Such a system will be more open to inspection and legitimate use by interested parties within the agency, while also requiring explicit data security policies, rather than the current practice of relying on security through obscurity.

2.7.5 System Reliability

A further difficulty in using data throughout an agency is concern with data quality. Even in systems where reliable, plentiful data are available, portions of the agency may not use the data due to skepticism in its accuracy. For example, this is the case at the MBTA, with high quality ATS data not used in planning in part due to skepticism of its accuracy. In some agencies the data remains in a fairly raw form, with substantial embedded “noise” or errors. Even in systems where vehicles are tracked imperfectly, it must be the case that the controllers are capable of using intuition and

logic to rule out impossible conditions and create a mental image of the state of a line. For a one time investment, this intuition and error correction can be embodied in algorithms to make the data useful to a wider audience. Given the decreasing constraints on data completeness and timing from operations to performance evaluation to planning, for systems with a control center where control is based on available data and minimal visual and radio information, this data should be sufficient for performance evaluation; if they are sufficient for supervision they should also be sufficient for planning functions. If agencies with centrally controlled operations lack the data to provide adequate control, then an effort to improve communications could lead to benefits across the agency.

2.7.6 International GPS Skepticism

GPS satellites are operated and maintained by the US Department of Defense, and are subject to shutdown, jamming and deliberate manipulation at the DoD's discretion. Russia has a competing system called Global Navigation Satellite System [Glonass] which has been plagued by funding difficulties.[24] It has not yet attained a full constellation of satellites needed for reliable positioning, but Russia has plans for a full constellation by 2005. The EU has been planning a third work-alike system named Galileo, which is expected to begin operation in 2008. High-end positioning receivers can improve their accuracy by using both GPS, Glonass and potentially Galileo signals, however the cost of integrated receivers are presently well beyond the reach of transit agencies. This skepticism has waned to some extent, and international GPS use may increase due to the promise of increased precision and security of Galileo, along with GPS interoperability.

Chapter 3

Method and Approach

This chapter defines a framework for the appraisal and fulfillment of the agency needs described in Chapter 2 using data available from existing and/or proposed automatic data collection systems.

This agency appraisal begins with a survey of agency sources of data. This includes all existing and proposed agency hardware currently providing, or expected to provide, information in a digital format. Next, this information on available data is combined with the categories of agency needs outlined in Section 2.1 and graphically presented in order to help focus attention on underutilized data sources. The appraisal continues with assessment of agency needs with potential to take advantage of agency sources of extensive data. For each agency need this section describes the current sources of data used in satisfying this need as well as the specific data sources needed for an improved level of performance possible by replacement or augmentation of the manual data collection procedures. The challenges of producing the specific data products needed for each task are explored in order to better understand the complexity of each possible application. Weighing project benefits in light of application complexity, urgency and cost, one or two applications are chosen for actual design and implementation.

The approach for identifying the most promising avenues for more intensive use of agency data is a systematic examination of: Agency data sources, agency functional needs, interdepartmental utilization of agency data sources and intermediate products. This includes the expected performance benefits available by improved uti-

lization of existing and proposed information sources. In order to narrow the field of choices, next it addresses the challenges of each application in terms of difficulty of generation, combination, distribution, and display of the information. Lastly one or two applications are chosen for implementation based on a ranking of benefits, difficulty, and time-frame. This selection is made in light of the value of the intermediate data products created which can be shared within the organization.

3.1 Agency Data Sources

This step of the framework is a survey of agency hardware which may be used as a source of electronic information about vehicles and/or passengers. The automated data collection systems listed in Section 2.3 are examples of such systems currently in operation in a number of agencies, but systems not previously mentioned may be included if they meet the criteria of being able to provide automatically collected data to the agency. Data collection hardware under consideration or in the process of being acquired and installed should also be considered at this phase to plan to make best use of the systems once installed.

3.2 Data Source Utilization

Using the description of agency needs found in Section 2.1 this step identifies all agency areas not currently being aided by automatically collected data. Through a quick graphical evaluation of the technologies available to an agency, compared with those agency data uses identified previously, one can identify areas to focus on. This is shown in the two part Figure 3-1. The figure is split into a portion for (near) Real-Time applications, and a portion for Post Processing. The dark lines in the figure indicate the primary agency use of the listed technology while the lighter lines indicate possible additional uses of the technology. Application to a given agency would eliminate the hardware which is not available, and identify the uses to which that automatically collected data is currently being put. The remaining “alternative

data use” lines are areas to consider more closely when considering uses of existing data.

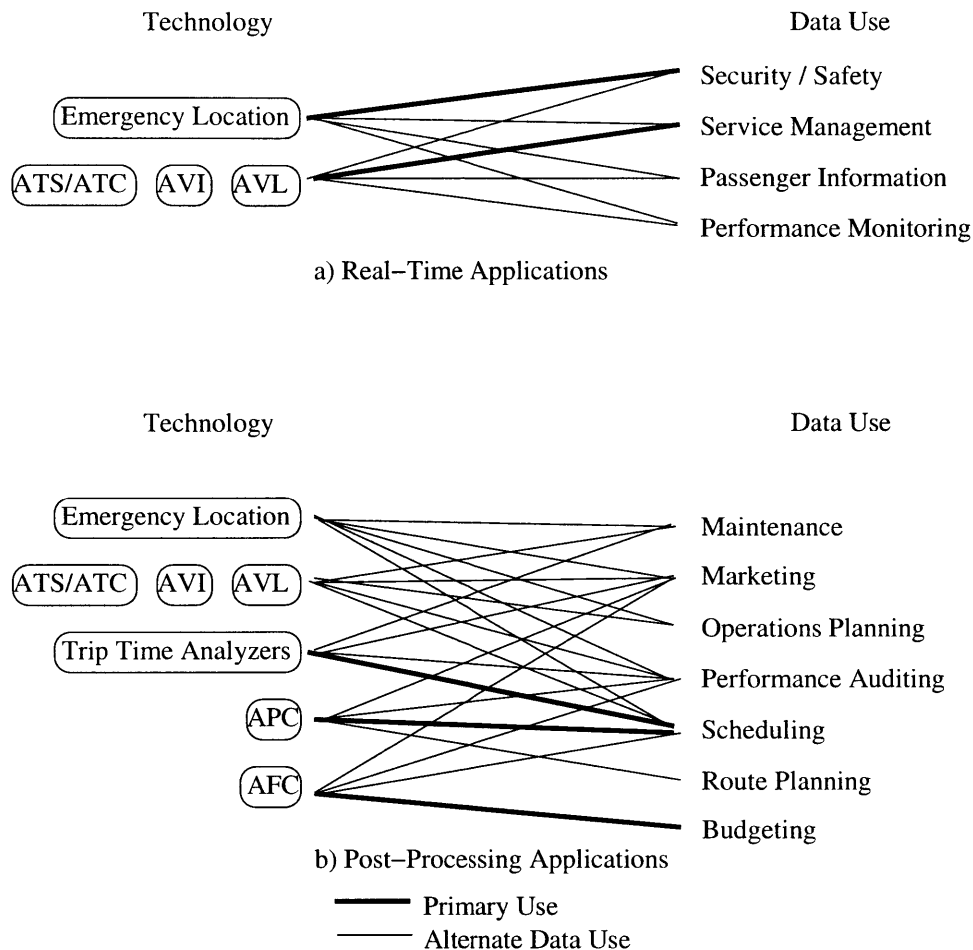


Figure 3-1: Mapping from Data Collection System to Agency Use

3.3 Agency Needs and Opportunities

In this step each of the agency needs described in Section 2.1, are addressed specifically: security, service management, passenger information, performance monitoring, maintenance, marketing, operations analysis, performance auditing, scheduling, route planning, and budgeting. For each of these needs, first the current agency data sources which feed that agency need are explored. For example to support the scheduling function annual ride checks for every route may be performed to provide

run time and load profile information. Next it attempts to identify an improved level of performance for that agency function, using data from a collection system listed in Section 3.1. For example in scheduling, using AVL derived position data to improve confidence in the setting of running and recovery times to meet agency on-time performance goals. At this point it also identifies this particular application with a unique identifier for later reference. Included in this step is the development of practical questions the agency can answer using this new source of data. For example in the scheduling case, the agency may be able to answer other questions about terminal performance such as: what is the minimum recovery time for an operator at the end of a trip? Lastly it identifies the specific data products needed to provide this improvement. In the case of the previous example this would require extensive data on run times in order to provide enough information to be confident that the required percentage of vehicles can start their trips on time.

3.4 Intermediate Data Products

To better understand the complexity of each possible application identified in the previous steps, this step summarizes the information needs of each of the agency applications identified in the previous section. From this it identifies the most commonly used data products, and provides information on the steps to derive the needed information from the available data sources. This step presents potential sources of individual data products needed for improved agency satisfaction of the functional needs listed in Section 3.3. It considers both data products currently generated by agency software tools, as well as products which would require new processing to create. The inter-relations of these intermediate data products should also be considered, as products created for one task may greatly simplify another separate task. An example of this is that a passenger information task which makes available processed AVL data to provide vehicle position along a route can then be readily used for performance auditing.

Figure 3-2 summarizes the inter-relations among intermediate data products. In

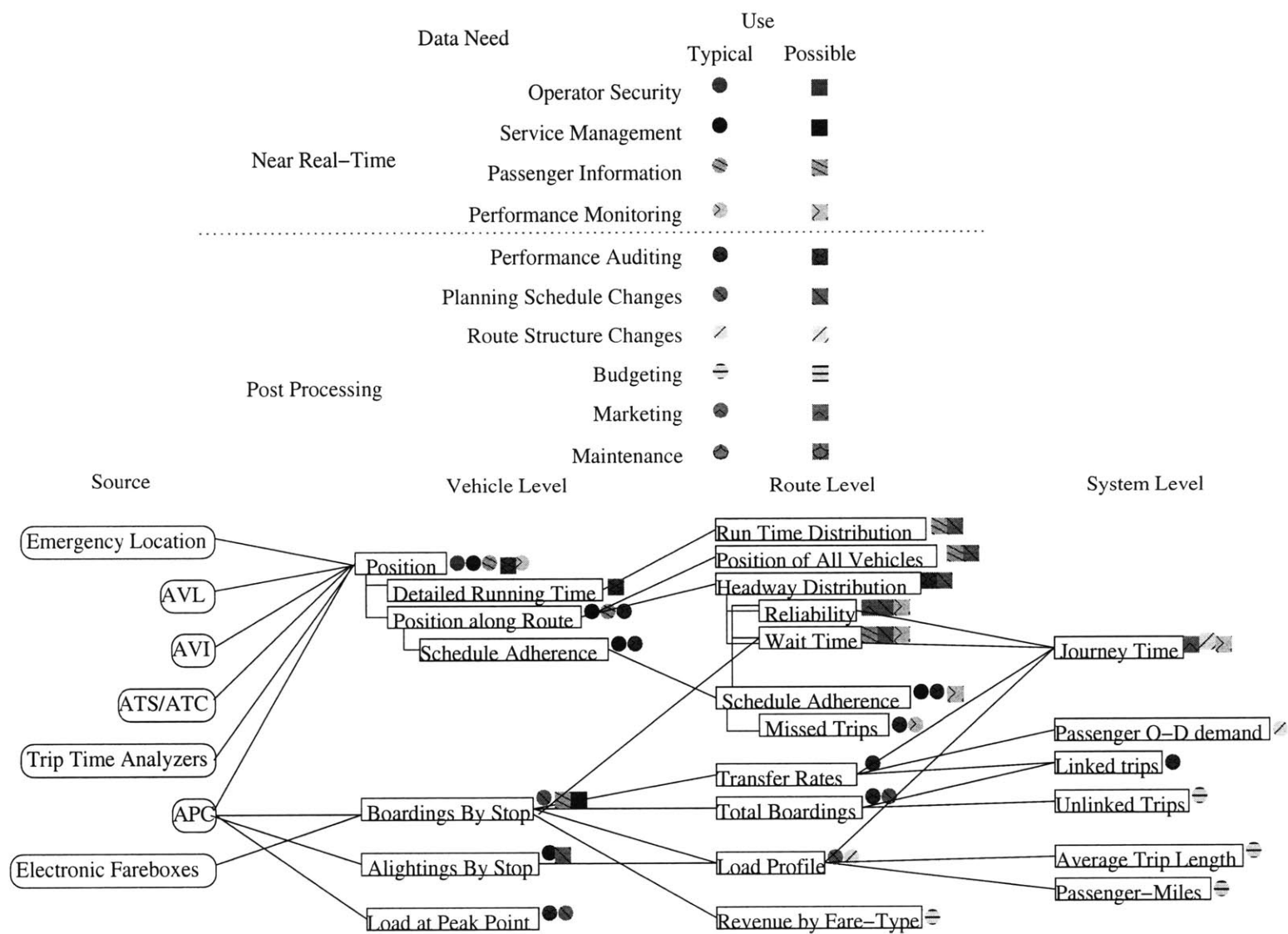


Figure 3-2: Summary of Agency Data Needs

this figure, data collection technologies are listed as feeding data products at the individual vehicle, route and system-wide levels of aggregation. Route-level aggregate figures are calculated using all vehicles on a route, or the fraction which are available. Certain data products, such as headway distribution and route level average wait time figures which are dependent on having sequential vehicles, will not “fail well”: a single missing vehicle will invalidate more than one instance of this measure, for example two headways will be unavailable if data is not available for a single vehicle. Other data products, such as schedule adherence and run time distributions are more robust with respect to data completeness at the route level. System-wide aggregate values combine multiple routes worth of readings to provide a single value to express system effectiveness or efficiency. Connections between data products, for example between “Position Along Route” and “Schedule Adherence” indicate that one is derived from the other using software processing. Agencies can use this figure both to identify particular sources for the data requirements to satisfy functional needs as well as to identify high-value intermediate data products that will be helpful to the largest number of agency tasks.

3.4.1 Application Challenges

This section describes the range of challenges presented by each application, in order to assess the time, difficulty and risk associated with each project. The ease or difficulty of generating the data in the required form for an application is important in planning implementation. This information helps management assess whether the in-house staff are suited to undertake a particular task, or if this task is best left for contract work. This section addresses each application identified in Section 3.3, exploring the range of challenges which need to be addressed including the programming of complicated algorithms, any challenges of information distribution, and unusually large quantities of data required. It also references the previous Section’s exploration of the particular data products used, assessing the difficulty of the creation of each.

3.5 Application Selection

In this step we decide which applications should receive priority in implementation within the agency. The selection process weighs expected benefits to the agency against application development time and cost. The expected benefits were addressed in Section 3.3 and the development time and cost as related to application complexity were discussed in Section 3.4.1. The degree to which existing manual data collection is failing to provide adequate information for effectively meeting the need increases the expected benefit from such a project. Certain projects may even have future cost savings as a side benefit of implementation, such as through the reduction of manual data collection and its associated labor costs. The complexity of the task will influence the time it takes for implementation.

There are a number of costs related to the project complexity, such as costs incurred in the creation of required custom software, as well as hardware costs associated with any needed improvements in communications bandwidth, data processing resources and data storage. Projects for which in-house staff is qualified and available to undertake the software challenge will be much more attractive than projects which require extra staff or the use of contract labor. Also in some circumstances underutilized hardware may be available at no cost; this will tend to favor projects which can use these resources. For a given project complexity, those which have a number of incremental sub-parts which are useful on their own are less risky than projects which are all-or-nothing in nature.

The central challenge of application selection is to bring together sufficient information on predicted benefits as well as sufficient information on costs in a way that they can be directly compared. This can be especially difficult as those with the technical expertise to understand each project's complexity may not have the transit expertise to understand each project's benefits. This is further complicated by likelihood of synergies between projects due to use of common intermediate data products. There are a number of ways of distilling cost and benefit information for use in decision making, but only two are addressed here: comparing relative costs and benefits

of projects, and comparing detailed software and hardware cost estimates along with benefit estimates. In the former, estimates are made of relative costs of a range of projects, as compared to the relative importance to the agency, and within a given time frame. For example a technical committee can rank the relative complexities amongst a set of projects and separate them into categories – ranking each on a scale of 1 to 5 for example – providing a rough estimate of when the required hardware will be in place to provide this data if it is not already. In parallel an executive or a committee from a range of departments ranks agency benefits from different projects.

A more careful assessment of project costs may be in order, if higher-risk large scale projects involving new hardware purchases and contract-written software are under consideration. One issue here is that an agency may not have staff skilled in making accurate estimates of the costs associated with large hardware and software projects. It may be that a small-scale, low-investment project is necessary to provide the information required to assess the requirements of a related larger-scale project. A more careful assessment will consider each component of cost in turn, including costs for new data processing hardware, new storage and archiving required, upgraded communications links and finally for the custom software to provide the integration. Even if this in-depth cost assessment is not done for each potential project, it may be used to assure that any tentatively selected project or projects are feasible given the intended time frame and capital budget.

First approximate processing requirements are estimated based on the expected algorithmic complexity combined with the input and output communications requirements. The archival needs of the agency users are anticipated with a similar process to determine required storage space given input data rates and anticipated retention requirements. Next determine a data flow path for each of the inputs and each of the outputs. In some cases this may be trivial, as when the data of interest are changed infrequently and are in a database – the data flow path is merely a remote database query over a local area network. In other cases this may be the most restrictive portion of the project, as when communication must take place over a shared radio uplink. In some cases the difficulties in this step may be deliberate, as in the case

when there is a firewall separating an internal system from an external system, with strict regulations on access across this boundary. Lastly, from the complexity of the data transformations required, determine whether the agency has the capability to write the software in-house in the expected project time-frame.

Both the in-depth cost method and the cost prioritization method will tend to overlook synergies between projects. One method of counteracting this is to use Figure 3-2 to identify projects which require the creation of intermediate data products of use to a number of other projects. In the ranking of benefit to the agency, the extent to which other projects are simplified should improve the rank of a keystone project, especially when a number of other projects are greatly simplified and when these projects provide vital agency benefits. Alternately, if one or two specific data products are identified as useful in a range of applications, then a separate cost / benefit comparison can be done conditioned on each of the separate products' availability. This may lead to a decision to go ahead with a particular task even if it is relatively complicated.

The prioritization of the improvements to agency functions, and what costs may be born for their implementation will be heavily influenced by agency concerns and values. In an agency with a strong customer focus, passenger information would be weighted heavily; in an agency with capacity and service quality concerns, then prioritization would tend to favor operations analysis and service management.

Chapter 4

Case 1: CTA Bus

This case study describes the application of the framework developed in Chapter 3 to the bus division of the CTA, with the addition of a section on actual agency benefits, and the lessons learned. It begins by appraising the agency's current ability to meet a range of needs, and identifying two promising software applications. One application uses operations data as a source for real time passenger information and service management. The other uses archived operations data for operations analysis. Next is a description of each selected application along with examples of improved agency performance using each application. Finally it closes with a discussion of critical decisions addressed in the implementation, including lessons learned in the design and implementation of each application to aid in future projects.

The CTA is the second largest public transportation system in the United States. The Bus division of the agency serves approximately 1 million passenger trips per day using 1900 buses on the system's 139 routes, covering more than 2,000 route miles on Chicago's predominantly grid street network. See Figure4-1. It serves more than fifty routes operating more than 18 hours per day. CTA has more than 11,000 employees to provide and manage this service.

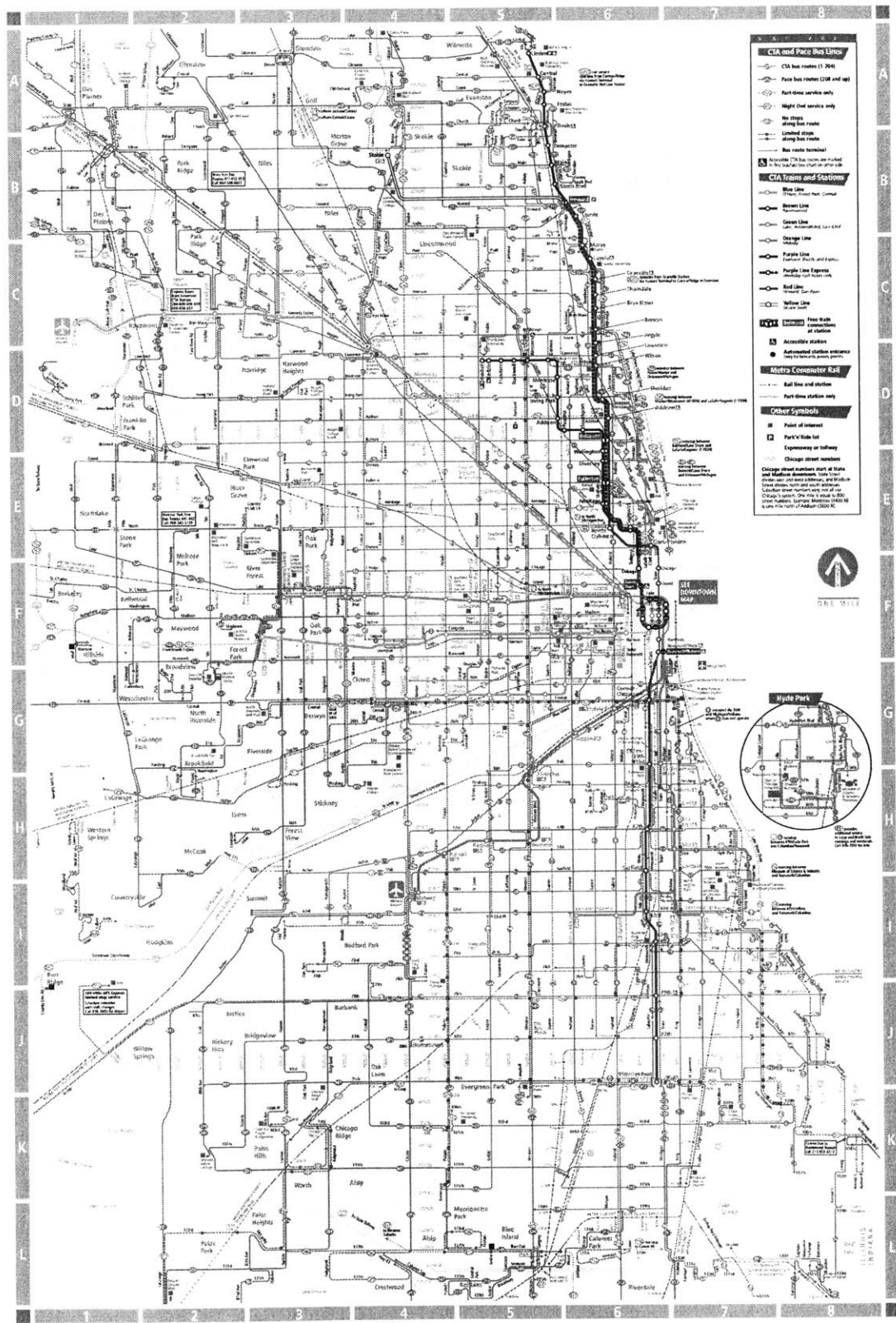


Figure 4-1: CTA System Map

4.1 Agency Appraisal

This section first reviews the data collection hardware currently in use in the agency, as well as hardware under consideration or in procurement. Next the available hardware is compared with the current uses of data sources in a graphical format to help identify especially promising areas for helping the agency meet its functional needs. Then comes an appraisal of the agency's ability to meet a range of needs, focused on those needs with potential for improvement through increased use of automatically collected data, from the short-term service management decisions in operations to the long-term strategic planning decisions of where to add, remove or reroute bus service. Specific options for improving agency fulfillment of each of the listed needs using automatically collected data are identified. The challenges presented by the generation of needed intermediate data products are presented next, followed by the challenges of each application, in turn. Based on knowledge of agency values and discussions with management, two modest applications with potential to benefit the agency emerge.

4.1.1 Agency Data Sources

A number of sources of data are currently or soon-to-be available to the CTA Bus division. Presently, an Emergency Location AVL system is available on 1600 vehicles to provide real-time position updates to the CTA control center. A transactional AFC system is available on all 1900 vehicles which records time of fare collection, but this is not currently linked to the position reporting hardware. Delivery has begun on a stop announcement system to be installed on approximately 1400 vehicles by the end of 2003 which should be fully operational by the end of 2004. Approximately 14% of the fleet, 266 vehicles, will be equipped with APCs for automated measurement of ridership.

The only data source which is available in near real-time is the Emergency Location system from Orbital Sciences. This system is currently justified by its safety benefits since it provides the ability to locate a desired vehicle and dispatch emergency

equipment without the operator having to identify and transmit the vehicle position verbally. The on-vehicle component of the Emergency Location system is the Mobile Data Terminal [MDT], which is part of the CTA's Bus Emergency Communications System [BECS]. Each MDT has a GPS receiver for vehicle positioning, a handset for voice communications and a microphone for covert monitoring of on-board audio when a silent alarm is triggered. The MDT acts as a digital interface to a three channel analog radio provided by a different vendor, and also has a digital interface to the vehicle odometer, although this input is not used in position reporting. The MDT also collects data about the operator; operators log in to their data terminal when they begin their shift, and log out at the end. To log in, an operator must key in the shift's unique identifier – the run number – as well as his or her unique badge number.

BECS was specified and delivered as an emergency communications system with its primary goal being to improve operator security through the silent alarm and vehicle tracking features. A subset [approximately 300] of these vehicles have an additional short-range radio installed along with upgraded software to provide an on-vehicle service management function, indicating minutes ahead of or behind schedule. This system, called the Bus Service Management System [BSMS] is not currently in use, due in part to difficulties with providing appropriately formatted route data to the MDTs. In order for BSMS to function properly, the vehicles all have to be based at one garage, but since the BSMS hardware was installed only on the subset of the fleet which is in the best condition, it has proved difficult to keep BSMS vehicles in a single garage.

The GPS derived position is available to control center staff in near real time. Historically this has only been available when an individual vehicle is queried for its position, however a recent modification to the MDT software transmits vehicle position to the control center for all vehicles with the new software. For off-line use, both data collected from the Emergency Location system as well as the transactional AFC data are promising resources for further use. The Emergency Location system has one additional feature for off-line use – a floppy disk drive that can be used to

record vital vehicle information at a rate of one record every four seconds. This data is very detailed, including information about whether doors are open, and odometer information. Unfortunately, the process required to collect the data is quite labor intensive, which negates some of the benefits of utilizing automatic data collection systems.

The AFC system is another on-vehicle system suitable for off-line use. It is comprised of on-vehicle fareboxes, and probes in each garage which communicate with a central mainframe computer to record the card number used, the fare amount, fare type [card fare, cash, discounted transfer, etc.] and the time to the nearest second. Unfortunately the AFC does not directly record either vehicle position or passenger alighting information and without this information it is difficult to infer more useful information such as a load profile for auditing maximum vehicle load. It may be possible to infer position from the sequence of AFC records, or to correlate Emergency Location and AFC data. It has historically also been difficult to extract large quantities of data from the AFC system. A number of standard products are available to analysts on their desktop, but accessing other data processing systems requires familiarity with the legacy systems involved.

The stop announcement system will also be a source of high quality recorded data. Clever Devices, Inc., the system vendor, refers to the system as an Automated Voice Announcement System [AVAS]. This system follows a “smart bus” design [see Section 2.5] where on board devices are integrated over an SAE J1708/J1587 communications network. System devices include a GPS and dead-reckoning enabled location system, a stop annunciator and destination sign unit, an operator console for sign-in and operator messages. It will provide an integrated, single login for the operator. This is in contrast to the present operator log-in which requires the operator enter badge number and run number into the MDT, route number into the fare box, and route variant number into the external route destination message sign. This single login will both simplify the operator experience and improve data quality. While the AVAS system does not provide data over the radio directly, it does communicate with the MDT, providing badge and run number, which is then sent over the MDT radio

data channel. When fully implemented, this system will have a database of CTA bus routes and stops and will identify every stop for passengers using an on-board variable message sign. It will also be capable of audibly announcing every stop along a route. The volume of the announcement will vary depending on the background noise level in the coach, in order to assure announcements are audible but not overwhelming. The system also announces the route name and destination each time the doors open. [16]

The system will record a number of events relating to vehicle progress along the route, each event will have a time-stamp as well as a location-stamp, derived from the dead-reckoning enhanced GPS. An event is generated when a door is opened or closed, as well as when the vehicle has a change in speed or heading, or a change in the operator sign-in status. Data will be offloaded nightly in the garage and transferred to a central Oracle database server. A number of systems are already installed and providing destination announcements. By the end of 2003, 1432 vehicles or approximately 75% of the present fleet, will have the hardware on board, and the system is to be fully operational by the end of 2004. All new vehicles purchased will have new AVAS installed, in order to bring the percentage of equipped vehicles in the fleet to 100% as vehicles are retired and replaced. [27]

A portion of the fleet will also be equipped with APC hardware. This component will be attached to the vehicle network, and generate a message indicating total ons and offs each time a door closes. The message will be logged along with other status messages for later offload and transfer to a central database. Approximately 14% of the fleet, 266 vehicles, will be equipped with APCs; this component will also be included in all new vehicle purchases.

4.1.2 Data Source Utilization

As illustrated in Figure 4-2, there is substantial opportunity for use of data in the overlap between agency needs identified in the appraisal and agency data collection systems. As discussed in Section 4.1.1, the bus division of the agency currently has an Emergency Location system as well as an AFC system, and will soon have a APC

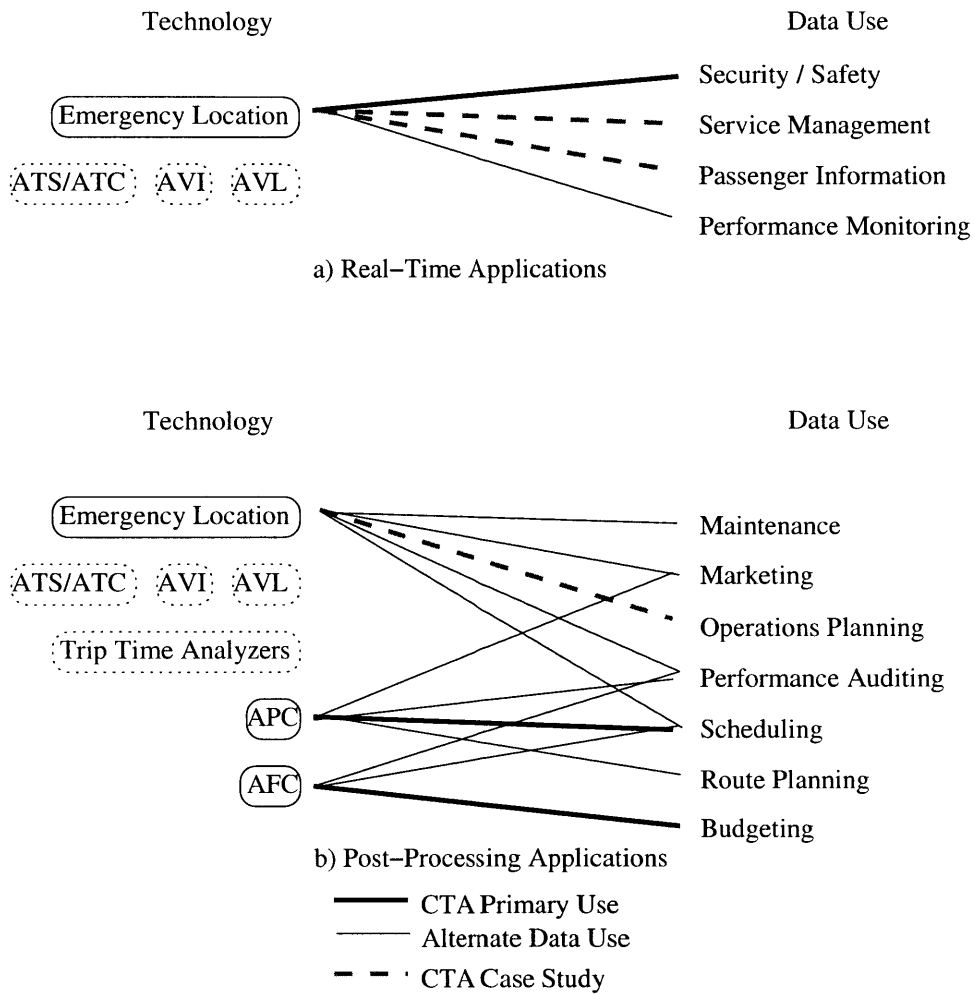


Figure 4-2: Case 1: CTA Bus Use of Agency Technology

system to utilize, so focus can be quickly placed on uses for these data sources.

4.1.3 Agency Needs and Opportunities

For each of the needs outlined in Section 2.1 this section identifies the degree to which these needs are currently satisfied and identifies potential improvements in performance possible given the particular sources of extensive data currently available or soon to be available to the agency.

Security / Safety

The security of passengers and operators is of great concern to the CTA. In 1999 a video recording system was installed on 1400 vehicles to deter incidents and aid in apprehension and prosecution of offenders. In addition to the deterrence factor and the effect on perceptions of security, surveillance camera images have so far led to 34 arrests.[1] As described in Section 4.1.1, the CTA has a communication and emergency location system designed to improve operator safety by providing a silent alarm function to covertly summon assistance. This system is installed on 1600 vehicles in the fleet, however the proper functioning of these devices is not routinely checked.

The combination of extensive data from both the Emergency Location system and the new AVAS system will provide an opportunity to identify and repair malfunctioning systems. This, application “sec1” has the potential to increase operator security by ensuring the MDTs will be able to reliably serve their function in an emergency. The combination of Emergency Location data with GIS route data in real time could provide automatic notification of an off-route vehicle, the possible victim of hijacking, however there is a substantial concern about “noise” in the position reports available from the MDTs, and this may generate too many false alarms to be of use.

Service Management

The agency currently uses controllers, as well as mobile and fixed supervisors for service management. There are 6 controllers during the peak period, and approximately 80 field [fixed and mobile] supervisors. Field supervisors can communicate with one another and with the controllers, and controllers alone can carry on radio communications with individual vehicle operators. Given the volume of urgent communications handled by the control center on any given day, fixed and mobile supervisors are limited to giving operators directions in person. For a discussion on the details of the extensive limitations of the present communications systems, see[2] and[25]. The CTA is also experimenting with an operator empowerment program to encourage operators on a set of routes to take measures, such as holding, expressing or even short turning buses to improve service without supervisor or controller intervention. The CTA has substantial problems with vehicle bunching, where a long wait at a bus stop ends in the arrival of two or more vehicles at nearly the same time. Street supervisors do not always hold vehicles to schedule, and there are often only one or two street supervisors along a route, not enough to guarantee reliable service. Not all terminals have fixed supervisors to assure on-time departures.

Data on vehicle position of all vehicles on a route or set of routes as provided by the Emergency Location system can allow controllers to much more effectively monitor service quality and intervene in cases where holding or short-turning buses will provide more even headways. This is application “sm1”. As discussed in Section 2.1.2 improved reliability of service can yield a number of benefits in cooperation with service planning: the CTA has the opportunity to provide an improved level of service on present routes, or expand service to new markets. An expanded version of this tool may included real-time load information in the future, though no source is available now. If load information were available to the control center in real time controllers could benefit from automated loading information by being able to quickly identify lightly loaded vehicles which are candidates for holding. Passengers would benefit from information informing them of the time to the next lightly loaded ve-

hicle's arrival. Data on vehicle position and load of all vehicles on a route or set of routes could allow controllers to much more effectively monitor service.

Dispatch control is vital to provision of reliable service; the ability to remotely monitor and direct vehicle departures has the potential to improve service quality by reducing initial headway variance. Dispatch control by finishing deployment of the the BSMS system, application "sm2", has the potential to greatly improve service quality.

Passenger Information

As mentioned above, the CTA bus network is large and complex covering nearly 2,196 route miles with 139 interconnected routes. High quality passenger information is an essential part of making such a system effective. The bus division currently provides users with a great deal of information, yet there are a number of options for improving fulfillment of the user needs discussed in Section 2.2. Passenger information needs are split up into pre-trip, at-stop and en-route needs. First is pre-trip planning; this is important to best utilize transfers between buses as well as to and from rail lines. These pre-trip needs are addressed using traditional paper maps, as well as web-based maps and a trip planner, and also a telephone information line. The Regional Transportation Authority's [RTA] web-based trip planner is accessible in only two clicks from the CTA website. The RTA telephone information line [312 836-7000] provides automated schedule information, as well as the ability to create an itinerary with a customer service representative seven days a week, 20 hours per day.

The web and phone services provide only schedule information, since they do not have access to real-time information on the actual state of service. Neither do they take historic measures of service variability into consideration when suggesting departure times and designing connections. This variability information may be easily collected using archived BECS data or AVAS data when available. For a bus to bus connection, information on the feeding line's ability to arrive at a point by a particular time, as well as the fed line's reliability at the transfer point could be combined to provide a more realistic estimate of expected transfer time, which may

be invaluable in making a transfer decision. A trip planning application [“pi1”] using static information which includes historical segment level run time distributions or simply variance information could support an “arrive at destination by a certain time” feature, which, given this variance information, could suggest connections with (say) a 95% confidence of being met. Experienced travelers likely have a sense of the reliability of their routes and transfers, however the trip planning tool could provide some element of this insider information even to new customers, suggesting how early to leave, and emphasizing higher reliability modes, including walking and rail. Neither real-time, nor historical detailed reliability information are currently available for CTA bus routes. Agency access to real-time vehicle information and detailed reliability information will be discussed in more detail in this Section’s discussion on Service Auditing and Monitoring. The difficulties of coordination in providing timely information to the RTA for a trip planner with real-time information, [application “pi2”] in a usable format will be discussed further in Section 4.1.5. Interface standards explored in Section 2.5 should minimize confusion.

The agency has 12,000 bus stops. A person waiting at a stop has a number of different options for information. At a few of the busiest stops schedule information and route maps are available for the lines passing by that point. Stops in the downtown area and many on busy routes have route maps with some information about span of service. Some stops merely have the CTA logo sign, without even mentioning the route name. Even these signs are useful, though, as they provide the phone number for the information line above. For users with a mobile phone, and for all users at stops near pay phones, all the trip-planning information mentioned above is available from the stop. No stops currently have real-time next bus arrival information, application “pi3”, although there are plans to provide such information to users in the near future in a RTA program with SAIC and NextBus.

On-vehicle information is presently limited to the destination sign and what is provided by the vehicle operator. Printed timetables are not typically available for distribution on buses. However this user information component will soon change as the CTA completes installation of the stop announcement system by Clever Devices.

Once fully operational, this system will provide announcement of the route, final destination, and individual stops.

Maintenance

Bus maintenance at the CTA benefits from some degree of automation. Trouble tickets created by controllers when a vehicle or vehicle component fails are entered into a maintenance database. Driver reports of items in less urgent need of repair may also be integrated in order to facilitate maintenance. There is an opportunity for more widespread, coordinated use of data from installed AVL and APC systems to ensure their health. Some of these systems fail in obvious ways – entering a fall-back voice communications mode, or failing to correctly debit swiped cards – and may be reported by drivers. Other modes of failure of on vehicle systems are more subtle – such as an AVL system providing erroneous position information. As the CTA AVL system is primarily for emergency vehicle location information, this problem may only be noticed when there is an emergency. Subtle problems may be symptoms of a larger problem, which could be identified and repaired before a more serious failure. Vehicle performance information could also be recorded and telemetered by information components, which may be correlated with increased breakdowns, such as hard acceleration, deceleration, or background noise levels. Use of data from on-vehicle information systems, application “maint1”, has the potential to reduce on-route failures.

Marketing

The majority of CTA marketing is targeted to regular riders. Some efforts are designed to inform riders of changes in scheduled service, such as extended Brown Line hours. Other marketing is dedicated to new services such as the ChicaGO card or the new 49X express bus service on Western Ave. Lastly, marketing has attempted to influence user behavior to improve cleanliness and reduce dwell times. Service for special events, such as the July 3 fireworks, is marketed to a wider audience through press releases. A new residents program mails a two-ride pass and transit map to residents with new

telephone service installed in the CTA service area.[3]

Application “mark1” explores how marketing may benefit from the use of figures derived from analysis system performance made possible through the use of automatically collected agency data. By using APC-derived dwell time data, marketing would be able to reduce overall dwell times and improve service quality by presenting a figure such as “Each rear-door exit saves everyone x seconds.” Forthcoming APC data in combination with survey data could provide information about the total-trip-time benefits of express service which may be presented to users, in order to encourage trying out the new service. Using AFC-derived ridership data a billboard such as “CTA bus service kept 125,000 cars off the road yesterday. Share the road.” Assuming current approximate weekday ridership number of 1 million riders, divided by four trips per person per day, divided by 2 persons per car yields 125,000 cars. Exploiting the fact that parking structures typically cost \$10,000 per space, by counting PM-peak weekday fare cards at AFCs in the loop stations, a sign which impresses upon the public the importance of the CTA could be: “If everyone drove downtown rather than taking the ‘L’, the new parking lots would cost \$X,XXX,XXX,XXX to build.”

Operations Analysis

Bus operations analysis at the CTA draws mainly from operations staff’s observations of current policies, and improvement through gradual refinement. Voice communications channels are presently in high demand, and the only direct channel between the supervisors on the street and drivers happens when meeting face to face. For more information on this topic, see [2]. The operator empowerment strategy discussed in the service management section for improving service quality was created from strategic thinking on ways to avoid communications bottlenecks in provision of reliable service.

Operations analysis can benefit from the sort of improved information on vehicles described elsewhere in the portion of this Section on Service Management. Extensive data can provide staff with an improved understanding of operations dynamics – the specific precursors of vehicle bunching on a given route, for example. Other precursors to unreliable service can be found through use of extensive operations

data, as application “anal1”, for example the existence of inter-operator variability can be assessed and reduced. Variability in behavior between operators can lead to small irregularities in service, which if not eliminated quickly, can progress into long gaps in service and vehicle bunching. The distribution of passengers along a given route – the load profile – is not typically known on CTA bus routes. This load profile can be created from APC data [application “anal2”] to provide a basis for changing supervisor positions and even revising their duties, as well as be valuable in operations, to provide an understanding of the expected impacts of policies on interventions such as short-turning and expressing.

Performance Monitoring and Auditing

CTA service performance monitoring and auditing is formally the responsibility of stationary supervisors who presently manually log vehicle arrivals, and bus controllers and mobile supervisors. The only current automated tool to support this function is a control center summary display of bus runs held in due to equipment or operator availability problems. This helps controllers decide whether to shift a vehicle or an operator from another route to provide better service, however it does not provide detailed information on operations. Information from supervisor logs could be used for after-the-fact performance auditing, however there are a number of problems with using supervisor log information to calculate detailed measures of line reliability, including:

- Not all the information noted by the supervisors is keyed into a computer. The total number of on time and late vehicles for a particular time period is entered, but no information about actual headways or individual vehicle arrivals is routinely made available on line. Nothing is recorded about the magnitude of the disturbances – how late the late vehicles are.
- There are too few supervisors posted along any given route – typically two – to determine the service quality at all points along a route.
- As mentioned in Section 1.2, the reliability of this data is suspect. Aggressive

audits of this source of data are not routinely carried out, and since the data is collected by a supervisor partially responsible for the on-time performance of the route that he or she is auditing, there is an inherent conflict of interest.

To illustrate these concerns an independent audit of service on four high ridership [over 20,000 daily riders] CTA bus routes (the 20/Madison, 4/Cottage Grove, 3/King Drive, and 63/63rd Street) was carried out by volunteers from the Campaign for Better Transit who stationed themselves at points along each route and collected vehicle arrival time information.[14] This simple study calls into question the CTA's reported 90% on-time performance figure, derived from the supervisor logs. Other audits of performance are regularly done as part of the scheduling process, as described in the portion of this Section on Scheduling. Even these regular checks by the scheduling department are insufficient to monitor the effectiveness of interventions, or show when interventions would have been appropriate. One source of frequent, unbiased, audits of service performance is through CTA employees. Employees in bus operations are encouraged and expected to be users of the system [13] indeed the top system users are publicly acknowledged quarterly. While this is only an informal measure, it does provide regular observations of a number of routes.

Using position reports derived from BECS data, the actual performance of routes can be measured and provided to controllers to help identify routes in need of attention. [application "pmon1"] An automated reminder of the quality of service experienced by passengers – such as percentage of passengers waiting longer than a certain threshold – could help improve the performance of new controllers and supervisors. Simpler warnings, such as a notice of a developing long headway, could also be of use in training staff to provide the best service possible.

Post-processing of BECS data provides the opportunity to audit line performance over time. [application "paud1"] The percentage of vehicle-time spent at less than half the scheduled headway behind its leader is an appropriate one for the CTA bus division, given the tendency of vehicles to form pairs and bunches of larger numbers. As discussed earlier, it is possible to calculate this figure using data that may be incomplete. Using these figures, one may compare performance from one day to the

next, and review situations which address the effectiveness of the service restoration techniques applied and alternative techniques which might have been used. With the CTA's poor state of mobile communications, evaluating alternative techniques which take advantage of proposed communications improvements may help make the case for investment to improve the communications infrastructure.

Work is currently underway at the CTA to provide supervisors with personal digital assistants [PDAs], which are small, portable computers. This would solve a number of the present problems with using supervisors for performance auditing. In the demonstration version of the software, the supervisor does not manually write the time of arrival, but instead that is automatically added; this would make deception or simple errors much less likely. Additionally, since all the data would be digitized, it would be available after the fact for further analysis. Using the supervisor PDAs as a source of data for performance auditing [application "paud2"] still has the difficulty of providing data at only a small number of points along a route. Vehicle behavior at intermediate locations may be important for determining the severity of problems. In a situation with two supervisors along a route, if vehicles arrive at the first supervisor perfectly spaced, and then three in a bunch arrive at the second supervisor, the quality of service provided to intermediate stops is unknown. This source of data would serve as a good cross check for other sources of data.

Scheduling

Traditionally, the service planning department at the CTA has relied on a staff of traffic checkers to do point checks. The loading data from these checks are used to ensure that the peak load standard specified by the CTA's Service Standards [29, p.20] are being met. Service standards currently dictate specific headways for service based on passenger flow per half hour. The checking staff is assigned to investigate customer complaints, evaluate service changes, and audit current service frequency. Data automatically collected from on-board AFCs, in combination with current route schedule statistics, are used to rank productivity in terms of passengers per platform-hour. [30] Routes with less than 30 passengers per platform-hour are subject to reductions in

frequency of service, span of service, route change or wholesale elimination. The CTA bus scheduling department also has a contract with Transportation Management & Design, Inc. [TMD] of San Diego to review the running time allocated to routes and revise schedules accordingly. This contractor has its' own staff of point checkers to manually record times along the routes under scrutiny.

The agency has neither the in-house staff nor the funding to hire a contractor to manually check the run times on all routes; currently only a subset of routes identified as having run-time and recovery-time related reliability problems are under scrutiny. While routes with performance issues are routinely reevaluated, a number of schedules were written years and years ago for dramatically different traffic patterns. Automatically collected sources of extensive data, including archived data from the current BECS system, could provide some valuable information for making schedules more realistic. [application "sched1"] More reliable and higher-resolution sources, such as the forthcoming AVAS and APC systems, will allow the agency to rigorously set run times on all routes, and for all time periods as part of a CTA schedule efficiency review. [application "sched2"] Using extensive data, routes with too little or excessive recovery time are more likely to be identified. Using the run-time variance available from extensive data, a target percentage of vehicles available to depart the terminal on time can be developed, such as a goal of 97%. Using extensive data and automated scheduling software, the vehicle requirements of such a standard can be identified and the goal can be temporarily increased or decreased as appropriate to the current vehicle and operator base. Using such data the CTA may be able to determine a figure for the minimum recovery time an operator requires at the end of a trip, in order to leave on schedule reliably. Extensive data on passenger load will provide the ability to determine load profiles – to know what portions of CTA routes have the heaviest demand per half hour. Extensive data will also allow the rigorous evaluation of CTA bus service against its Service Standards, allowing review of every route, every season.

Route Planning

Under the direction of Dr. Michael Shiffer, the CTA is making great progress supporting route planning with a number of tools and data items that are part of its “spatial data infrastructure”. This includes census data, housing density, type of land use, aerial photos, route infrastructure, current schedules, the mix of vehicles available to the agency, as well as ridership data. There is currently a bus stop spacing program at the CTA which uses the GIS to highlight stops which are closer than the systemwide minimum of every standard Chicago block - 1/8 mile, and trigger further investigation. The service change committee is also using integrated sources of data to re-evaluate bus lines given present traffic generators, which may have changed dramatically since the route was originally designed. Examples of this from the North side of Chicago include restructuring express service to the loop into a zonal express service with four different zones. On the South side, express services which were all on Jeffrey St have been spread to collect passengers from three parallel streets, based on express bus demand. [6] One area where extensive data can facilitate this effort is with application “plan1” the estimation of origin-destination [OD] demand profiles from APC data. This OD information may be used as part of an investigation as to whether zonal express service would be preferable. Census data on housing and employment density as well as APC-derived load profiles from existing routes can provide the information to predict which corridors may benefit from express service introduction.

Budgeting

Budgeting presently is concerned with data collection only to the extent that data on unlinked passenger trips and passenger-miles figures are collected with sufficient accuracy to maximize US Government funding received. While a few of these figures are currently automatically collected at the CTA, most are collected manually. Use of data automatically collected by the AVAS system has the potential to reduce operating expenses, and reduce the variability in operating expenses currently expe-

rienced. Schedules which do not take present traffic into account may both result in irregular service and cost the agency more money than budgeted in overtime pay. Using automatically collected data, it is possible to evaluate the economic impact of different levels of compliance with minimum service standards. When automatic data collection systems have proved their reliability, eventual reductions in staff dedicated to providing point checks can be expected. A reduction in the level of street supervision is another expected product of further use of automatically collected data, as on-time performance data need no longer be manually collected, and service management functions may be handled more efficiently by central controllers with better vehicle position data.

4.1.4 Intermediate Data Products

This section reviews agency sources of data products required to fulfill the needs listed in Section 4.1.3 by first summarizing each application's data needs, to the level of identification of the availability of the auxiliary data required to generate each listed product. Figure 3-2 provides a template for identifying the sources which can generate particular data products. It next identifies the most commonly used intermediate data products based on finding multiple instances of a product in the application data needs listing. Lastly, it provides detailed information on the generation of each commonly used data item, along with a discussion of the anticipated difficulties in this task.

Listing of Application Information Needs

Using Figure 3-2 as a guide, this subsection summarizes the information needed for each application in order to identify the most commonly used intermediate data products. Once an intermediate data product is found to be used in multiple applications, it is assigned an identifying letter. This letter is used later in the list to indicate that detailed information on how to generate that intermediate product is found in this section and if this application is implemented then those data products with letters

will simplify creation of other applications.

- Emergency location system health monitor [sec1]
 - Most recent BECS Position report from each vehicle
 - AVAS Position data from all vehicles
- Line management using BECS data [sm1]
 - Position for all vehicles on a route [P]
 - * Position along route for arbitrary vehicle.
 - Vehicle coordinates: Latitude and Longitude from BECS.
 - Route structure: Ordered list of bus stop coordinates.
 - * Run numbers of vehicles of interest.
 - Vehicle number to run number mapping from BECS.
 - * List of vehicles on a given route.
 - Run numbers on a given route: from schedule.
 - Mapping of run numbers to vehicle numbers: from BECS.
 - Schedule adherence [S]
 - * Schedule information
 - * History of position of all vehicles on route [H]
 - Position for all vehicles on a route [P]
- Dispatch control using BSMS [sm2]
 - Reformat schedule and stop data for BSMS vehicles.
 - * Detailed schedule data from Hastus.
 - * Stop position list from spatial information database.
- Trip-planner with reliability info [pi1]
 - Segment-level run-time / reliability [R]
 - * History of position of all vehicles on route [H]
 - Schedule adherence [S]
- Trip-planner with real-time information [pi2]

- Position for all vehicles on a route [P]
 - Vehicle progress predictions [I]
- Real-time information for next-bus signs [pi3]
 - Time until next bus arrival at a stop
 - * Position for all vehicles on a route [P]
 - * Vehicle progress predictions [I]
- Maintenance use of vehicle performance data [maint1]
 - History of position of all vehicles [H]
- Using system performance in marketing [mark1]
 - APC card-dip entries
 - AFC demand records
- Identify precursors to unreliable service [anal1]
 - History of position of all vehicles [H]
 - Operator login information
- Load profiles based on APC data [anal2]
 - Load profiles / demand info [L]
 - * APC position history
 - * Surveys to calibrate accuracy of assignment process
- Performance auditing using BECS data [paud1]
 - History of position of all vehicles on route [H]
 - (optional) Load profiles / demand info [L]
- BECS data for performance monitoring [pmon1]
 - Real-time position of all vehicles on route [P]
- BECS / Demand data for performance auditing [paud2]
 - Real-time position of all vehicles on route [P]
 - Load profiles / demand info [L]

- Transfer information for passengers.
 - * AFC / fare card system.
- Performance monitoring using supervisor PDA data [pmon2]
 - Supervisor PDA time-at-position data
- Archived BECS data in scheduling [sched1]
 - Segment-level run-time / reliability from [R]
- Archived AVAS data in scheduling [sched2]
 - Segment-level run-time / reliability from [R]
- Cross-check APC and AFC data [sched3]
 - APC position history
 - AFC records fused with position
 - * Transactional AFC records
 - * History of position of all vehicles [H]
- Estimation of OD matrix from APC data [plan1]
 - APC position history

Application Information Needs Summary

These letters are also used in Table 4.1 which summarizes the dependence of the listed applications on the four intermediate data products identified with a letter in braces in the preceding listing of application information needs. Applications which use [R] also use [H] indirectly, and this is made explicit in the above table, with a lowercase [h], to emphasize the importance of the underlying data source, as with users of history data [H] which is derived from real-time position data [R], in this case denoted as the lowercase: [r].

Intermediate Data Products Details

This section addresses the challenges in producing the above-listed four shared intermediate data products. The full list of these shared products is found in the key to

Application Description [ID]	P	H	R	L	S	I
Emergency location system health monitor [sec1]						
Line management usint BECS data [sm1]	P				S	
Dispatch control using BSMS data [sm2]						
Trip-planner with reliability info [pi1]	p	h	R	L	S	
Trip-planner with real-time information [pi2]	P	h				I
Real-time information for next-bus signs [pi3]	P	h				I
Maintenance use of vehicle performance data [maint1]	p	H				
Using system performance in marketing [mark1]				L		
Identify precursors to unreliable service [anal1]	p	H				
Load profiles based on APC data [anal2]				L		
Performance auditing using BECS data [paud1]	p	H		L	S	
BECS data for performance montioring [pmon1]	P					
BECS / Demand data for performance auditing [paud2]	P			L		
Performance monitoring using supervisor PDA data [pmon2]	P					
Archived BECS data in scheduling [sched1]	p	h	R			
Archived AVAS data in scheduling [sched2]	p	h	R			
Cross-check APC and AFC data [sched3]	p	H				
Estimation of OD matrix from APC data [plan1]						

- P Real-time position of all vehicles on route
- p Real-time position of all vehicles on route needed indirectly
- H History of position of all vehicles on route
- h History of position of all vehicles on route needed indirectly
- R Segment-level run-time / reliability
- L Load profiles / demand info
- S Schedule adherence
- I Vehicle progress predictions

Table 4.1: CTA Application Shared Information Needs Summary

Table 4.1.

Position for all vehicles on a route [P] Most needs within the agency have some use for extensive data on vehicle position, either directly or in derivative form. Position data is now available in real time only from the Emergency Location system, while higher quality position data will become available off-line from the AVAS system currently being installed.

Position of all vehicles on a route is a data product of broad appeal. It can be used directly for service management and passenger information, or with further processing for a number of other tasks, described later. This data product can be generated from data that can be gathered from the current Emergency Location system. The input data required to calculate each vehicle's position along the route include the current vehicle position in latitude and longitude and some information about route structure, in order to determine how far along the route a given vehicle is. Given appropriate software, the vehicle latitude and longitude can be obtained from the Emergency Location system, and in the near future, this information will be available without special access methods. The route structure, which can be treated as static, is available through the ordered list of the coordinates for bus stops along a route. A two step process is also needed for the determination of the vehicles on a route, as BECS does not explicitly keep track of which route a vehicle is serving. While the operator does enter the route number into the vehicle farebox and destination sign, this information is not available in real-time over the radio via the MDT. The operator does enter a run number into the MDT, which is associated with a route using the current schedule. In the relatively rare case when one vehicle starts off on one route and finishes on a different route – interlining – a given run number is associated with two different routes. The list of all run numbers on a given route is also needed, and this is available from the schedule.

As discussed in Section 4.1.1, once the AVAS system is on-line it will provide a single, integrated, login for the operator. On routes with 100% AVAS vehicles, it would greatly simplify this step if the AVAS route number was available over the

radio data link in addition to the run number as is currently designed.

Position for all vehicles on a route: [P]

- Position along route for arbitrary vehicle.
 - Vehicle coordinates: Latitude and Longitude from BECS.
 - Route structure: Ordered list of bus stop coordinates.
- Run numbers of vehicles of interest.
 - Vehicle number to run number mapping from BECS.
- List of vehicles on a given route.
 - Run numbers on a given route: from schedule.
 - Mapping of run numbers to vehicle numbers: from BECS.

History of position of all vehicles on route [H] In this case, collecting data for a position history record of all vehicles on the route is equal in difficulty to providing real-time position information and simultaneously logging the data. The only challenge here is making certain not to fill up the device that's being written to. If the data is being written to a database then the database storage as well as the index storage may potentially run out.

Segment-level run-time / reliability [R] Extensive vehicle positions recorded by the BECS system should be sufficient for the majority of the agency's segment level run time data needs. However, due to GPS shadowing, BECS data is not reliable downtown and since many routes have a terminal downtown, this is an important flaw. It could, nevertheless be used to augment manually collected data, and provide a great deal of information. The high quality data that will be collected by the AVAS system, and especially the APC system, will provide a great deal more information about run time distribution. Hastus, the scheduling package used by the CTA has some ability to process position reports and revise the schedule accordingly but it may be preferable to have this set of segment-level run-times calculated outside Hastus in order to assure this data is in an accessible format for use to support other agency

needs. One such use is the real-time vehicle position predictions used for next-bus information, described previously. As discussed in Section 4.1.3, the figures on run-time variance available from extensive data collection allow run times and recovery times to be set based on desired percentages of vehicles available to start trips on time. Suitable reliability percentages must be developed for the CTA.

Service variability figures from the extensive run-time data used in schedule tuning can also be used for other tasks if the data is in an accessible format. Operations analysis applications using this data for repositioning of fixed supervisors would identify locations along routes where variability begins to develop, or to identify routes with more severe variability problems for a higher level of supervision. Another use would be to aid pre-trip passenger information, even without real-time passenger information, by taking service variability into consideration when suggesting departure times and designing connections. As discussed in Section 4.1.3, a trip planning tool with access to run time variance or distribution could provide passengers with a better estimate of when they should leave in order to get to their destination by a certain time.

Data items needed for segment-level run-times and variances: [R]

- Schedule adherence figures, as above, sufficient for several observations in each time period.
- Locations of time points
 - From spatial information database.
- Decisions on %-ile of trips that should be on-time according to schedule. In a system which disallows running early, this is a trade off between reliability and speed.
- Decisions on %-ile of buses that should be available to start the next trip on schedule.
- Decisions on data collection accuracy. See [40]

Vehicle progress predictions [I] There are a range of alternate mechanisms for predicting vehicle travel; four such mechanisms considered for the CTA are:

- The simplest prediction is that each vehicle on the route will travel at a fixed average speed. This is unrealistic in areas of the route that regularly deviate from this average speed, such as areas of frequent congestion or free-flow. However, it is very easy to implement.
 - Historical average speed, derived from the difference in time of terminal departure to time of terminal arrival, divided by the distance traveled. This may be as simple as the travel time of the previous vehicle over that segment, which takes current conditions into account, but is slightly more difficult to calculate.
- Predicting that each vehicle will continue to travel at its current average speed. The realism of this is dependent on the length of the time window used to compute average speed, but it has the advantage of being easy to implement.
 - Previous vehicle position reports.
- Using an average speed value for a given route segment during that time of day. This may provide a good balance between accuracy and ease of calculation. Such figures are not yet available for CTA routes, but could be developed to feed this and other agency applications. This has the added advantage of being able to provide a segment-level run time distribution, which is of interest for providing a %-ile prediction for a margin of safety.
 - Extensive historic segment-level running times. [R]
- Using a model of vehicle speed on the route given time of day and the street characteristics, such as traffic lights. This has the potential to provide accurate predictions, however it is likely to be difficult to calibrate.
 - Historical segment-level running times.
 - Geographic information, such as number of signalized intersections, and number of lanes.

Load profiles / demand info [L] There are a number of possible uses of extensive boarding and alighting data at the CTA. Extensive information about boardings is presently available from the fleet-wide AFCs. Some information about alightings may be inferred from linking passenger trips using AFC data, see [9], however this relies on people using fare-cards and making symmetrical trips. The APCs that are in the process of being installed on 14% of the fleet promise to provide much higher quality data on boardings and alightings.

The distribution of boarding and alighting demand along a route is valuable information. This data can be used on its own for operations analysis and planning – deciding where to short turn a bus, for example. It can also be synthesized into a load profile, showing where vehicles are most heavily loaded on a route. This will be readily available from the APCs since the APC contract includes the installation of RSM systems’ RideCheck Plus software, which will provide this data to the CTA automatically. This presents a number of opportunities for the CTA, including rigorous evaluation of CTA bus service against the Service Standards [29].

Load profiles / demand info: [L]

- APC position history
- Surveys to calibrate accuracy of assignment process

Schedule adherence [S] While there is some discussion of the use of headway adherence for service management at the CTA, schedule adherence is a more commonly used criterion. Headway adherence is also less robust in a system where data from some active vehicles may be missing. In order to more effectively use the position of all vehicles on a route [P] for service management, the position should be combined with the schedule information for the time period.

Data items needed for schedule adherence: [S]

- Route-wide vehicle position, as above. [H]
- Schedule information for any given position along a route.
 - Internal schedule information on “supervisor sheets”.

4.1.5 Application Challenges

This section reviews the challenges to implementation of each application identified in Section 4.1.3. This includes identification of difficult algorithms, challenging information distribution, unusually large quantities of data required, or special-purpose combinations of data sources. This section also references the difficulty in intermediate data product creation addressed in the preceding section. All of these factors give a sense for the application complexity, cost and time required to completion.

Emergency location system health monitor [sec1]

As mentioned previously, data from the MDTs and AVAS may be used to cross check one another. While entering a fall-back voice communications mode is an obvious error, other modes of failure of on vehicle systems are more subtle – such as the MDT providing an erroneous position. Discrepancies between archived MDT positions and AVAS positions can be automatically flagged and quickly investigated to determine which device is at fault. In order to do this a recent BECS Position report from each vehicle, as well as AVAS Position are required. The first check is to identify vehicles which are not providing data from both systems. A second check is for the percentage of time that the position reports agree. As the AVAS is expected to provide higher quality data than BECS, some disagreement is to be expected between the AVAS and BECS positions, however gross position disagreement for a large percentage of the time can be automatically flagged for further inspection. Such automated indications of faults in the emergency location system provide an opportunity to enhance operator safety through the prompt identification and repair of faulty systems so they may reliably serve their function in an emergency.

Line management using BECS data [sm1]

The two main challenges of line management using the presently available BECS data are the identification of position for all vehicles on a route, denoted as [P], above, as well as the automatic creation of schedule adherence [S] figures. Schedule adherence

can initially be done manually, by human comparison of times to vehicle schedules, but an automatic comparison should be an integral part of any line management tool. Visualization of this information provides controllers with the information needed to quickly assess the performance of a line as a whole, as well as identify operators violating rules by running ahead of schedule. The provision of the graphical display of these figures is straightforward given the advanced state of graphical software creation libraries. No current or proposed vehicle systems at the CTA can provide load information to the control center, though such systems are possible given current technology. In the CTA system, if the opportunity for a software change to provide load information over the radio link arises, it is worth investigation.

Dispatch control using BSMS [sm2]

Dispatch control is possible given data on vehicle position of higher quality than is currently available from the BECS. If some system, such as AVAS, were able to telemeter its high quality position data in real time, it might be possible to use this data to enable service management interventions to reduce initial headway variance. Since radio channel bandwidth is an issue, providing an on-vehicle system for this function would be ideal. Such a system already exists in the BSMS-equipped vehicle fleet. To make this useful would require combining the agency schedule and stop-databases into a BSMS format. This requires the integration of detailed schedule data from Hastus, also used in schedule adherence, with stop position lists for the routes in question from the spatial information database. The combination of these data sources is straightforward, however the largest technical issue is the difficulty in obtaining the component data items. There is still an issue with equity, in terms of running all the best vehicles on a handful of routes. Perhaps these routes could be rotated, to provide high quality service to different areas over time.

Trip-planner with reliability info [pi1]

Another use of both the segment-level run-time and reliability [R] data, as well as schedule adherence figures [S] is an improved pre-trip passenger information appli-

cation. Even without real-time passenger information, by taking service variability into consideration when suggesting departure times and designing connections, such a system has value. As discussed in Section 4.1.3, a trip planning tool with access to run time variance or distribution could provide passengers with a better estimates of when they should leave in order to get to their destination by a certain time. Adding the reliability calculations to the trip planner is non-trivial, and there is the added difficulty that the trip planner is under control of the RTA and uses proprietary software which will be difficult to change.

Trip-planner with real-time information [pi2]

For real-time passenger information, users making short-term pre-trip planning decisions need the expected arrival time of the next vehicle at their chosen stop. This is a derivative of the first mentioned product: position of all vehicles on a route [P] and also requires travel time prediction from current vehicle positions to future positions. [I] There are a wide range of possibilities spanning a range of difficulties for future position prediction. The combination of the data elements is not difficult, and presentation of this information can range from a full web-based trip planner to a simple on-line next bus count-down. This project has built-in flexibility.

Real-time information for next-bus signs [pi3]

Similar to the real-time trip planning application “pi2”, information for next-bus arrival signs requires real-time position for all vehicles on a route [P], which is fairly difficult to obtain, as well as predictions of future vehicle positions [I], which presents a range of difficulties. The provision of vehicle arrival information to next-bus arrival signs presents a real challenge; a flexible simple solution is a solar and battery powered unit with a cellular telephone data link to a central server. Fixed infrastructure or a small fraction of a dedicated radio channel would reduce operating expenses, but increase the the initial system cost.

Maintenance use of vehicle performance data [maint1]

A number of the previously listed combinations of data, including the MDT and AFC health checks, ease the burden on the maintenance department. Regular checks of on-board equipment which would otherwise be done on a per-bus basis can be done fleet-wide at a fraction of the cost. Another category of data combination is combining vehicle performance data [H] recorded by information components, such as the MDTs or the AVAS, to provide figures that may be correlated with increased breakdowns such as hard acceleration or deceleration, or possible vehicle background noise levels. The stop announcement system provides stop announcements at a louder volume when background noise levels increase. If this noise level can be recorded as part of the standard AVAS data set, noise levels may provide warning of impending failure, allowing a vehicle to be pulled out of service before failure on the street. These data items would be combined with the date and time of actual failure, to identify potential correlation between failures and each of the potential indicators. The actual identification of failure precursors is a difficult task most likely involving change over time and observations of failures.

Using system performance in marketing [mark1]

Novel marketing uses of currently available AFC data may be simple to provide as well as effective at getting a transit-positive message out. AFC data can currently provide data with one days lag on total number of passengers, which may be processed into a number of items such as a predicted number of auto commutes or gallons of fuel saved. More involved data analysis could estimate savings in the number of auto trips into a particular area, such as the loop on any given day. From that, the various costs of providing private auto service for those people can be shown to be a bargain. More extensive APC data has other potential applications in marketing. It is possible to calculate figures about operational characteristics, such as estimating a dwell time function from APC data. With this thorough understanding of line performance marketing could present numbers on the importance of leaving through the rear door

if people are waiting to board at the stop.

Information about new policies which reduce total trip time as inferred from vehicle position [H] and expected load [L]. Integration of these figures into an informational campaign about new express-bus service and marketing favorable travel times created by improved operations on the BSMS route, or other routes with BECS or AVAS instrumentation providing a log of position history. Calculation of predicted alighting demand are simple algebra when combined with load profiles and AFC observed boarding demand.

Identify precursors to unreliable service [anal1]

Synthesis of the passenger-centric figures derived for performance monitoring combined with explanatory information, for example operator id, can be used to identify operators who are consistently providing high-quality service as well as those who are in need of re-training. Specifically, a simple method for identifying operators who behave in a manner conducive to bunching is to combine extensive AFC data with operator identification data. In a system with readily accessible time-stamped AFC data, comparisons of total numbers of passengers served by operators in a particular period may provide an indication that some operators are consistently serving more than the period average while others are serving fewer than average.

Data for per-operator total passenger count figures:

- Operator badge number and vehicle number pair.
 - From emergency location MDT login.
 - From stop announcement system login.
- Boarding times for each vehicle on the route.
 - From AFC system, selecting boardings for a particular time window over a large number of days.

An alternate mechanism for operator evaluation is to use the operator id combined with MDT-derived position information to provide a distribution of run time for a

particular segment, for each operator. Specific operators with average run times significantly lower than average may be encouraged to slow down, and slower operators may be retrained.

Service variability figures from the extensive schedule data used in schedule tuning could also be used for other tasks, if the data is in an accessible format. An application of operations analysis using this data could move supervisors to locations along routes where variability begins to develop, or moving supervisors to routes with more severe variability problems.

Data for per-operator run-time figures:

- Operator badge number and vehicle number pair.
 - From emergency location MDT login.
 - From stop announcement system login.
- Position along route for all vehicles on routes of interest.
 - From emergency location system data.
 - From stop announcement system.

Load profiles based on APC data [anal2]

The creation of load profiles is an intermediate data product, [L] described in Section 4.1.4. This task just facilitates others.

Performance auditing using BECS data [paud1]

Another measure using extensive position data is to quantify the percentage of time a vehicle spent at less than half the scheduled headway behind its leader. This figure is simple in the sense that it only requires processing a single extensive data set of a route's worth of position data to quantify segments of running closer than a threshold, and compare that to the total number of running segments in the data set. This can be calculated from a history of position of all vehicles on route [H]. Gaps in the data where vehicles are running the route but not registering force utilization of degree of

bunching as an indirect measure of service reliability. A more direct approach where headway regularity is counted will not provide accurate numbers when data is absent.

To provide auditing over time, categorizing data from multiple days by time period as well as by weekday and weekend, and after eliminating holidays, meaningful comparisons may be made over time. Such comparisons can provide operations management with numerical feedback on the effectiveness of service restoration technique changes, as well as run and recovery time changes.

BECS data for performance monitoring [pmon1]

Another figure derived from position of vehicles along the route [P] is a real-time automated warning of sub-standard service, for example to notify supervisory staff of low quality service that should be investigated. This can be especially valuable as an aid in supervisor training. An automatic trigger on the schedule adherence calculated for service management could also be used for performance monitoring. For example, to notify supervisors of individual operators ahead of schedule at a time point. A notice of a developing long headway could provide a warning of degrading service that is more closely aligned with the passenger experience. With high frequency service the passenger is not bothered by a bus that is running early, whereas a gap in service will be noticed. Headway checking would require a high level of vehicle maintenance to assure that the majority of vehicles are properly identified as on-route, otherwise non-existent long gaps in service would be a distraction. In order to trigger on ahead-of-schedule vehicles schedule adherence figures are needed in addition to route-wide vehicle position [P]. “Paud1” and “Pmon1” accomplish the same processing, just in different time scales. The two applications have very similar complexity to one another.

BECS / Demand data for performance auditing [paud2]

Integration of vehicle position information from the entire fleet along with AFC data on transfers and the load profiles generated from APC data can provide an integrated, automated reminder to managers of the quality of service experienced by customers,

taking at stop, in vehicle and transfer times into account. Figures summarizing this information, such as the percentage of passengers waiting longer than a certain threshold – can help improve the performance of new controllers and supervisors. This improves upon “paud1” by combining position along with a passenger assignment function identifying transferring passengers as well as boarding and alighting demand [L] from the APCs or from the AFCs using estimated alighting demand. The added complexity is rewarded by a more representative summary of line performance from the customer’s point of view.

Performance monitoring using supervisor PDA data [pmon2]

Using supervisor PDA time-at-position data, line performance can be automatically tracked; the supervisors enter vehicle passings which are put in a control center database in real-time. This database can be queried to provide a report of vehicle headways to controllers as well as to all supervisors along the route, allowing supervisors to hold vehicles to reduce headway variance. This provides another redundancy source for on-board data collection, as both the emergency location system’s MDTs and the AVAS will provide positions. Supervisor PDA logs of vehicle arrivals will provide the second real-time and third overall record of vehicle time at particular points. Checks of MDT and AVAS data against supervisor PDA data can automatically identify faulty hardware as well as suspect data entries by supervisors. This provides an improved ability to audit the supervisors themselves. The actual comparisons are simple, after the steps have been taken to make the real-time position data available from the MDTs.

Archived BECS/AVAS data in scheduling [sched1/sched2]

Both applications sched1 and sched2 provide data useful to the agency in a number of different applications. The creation of the actual segment-level run times [R], from position history data [H] is straightforward as described in Section 4.1.4. The application of these values provides some challenge, both in formatting the calculated values to be understood by Hastus, the scheduling package used at the CTA, assuring

the data used is not biased in some way, and also in identifying desired performance goals for on time performance and recovery, as running time trades off trip-time for wait-time, and recovery time trades off possible overtime for longer scheduled hours. Extraction of archived BECS data to provide the needed position history is fairly difficult, and the data provided are not the best quality, due to over-air bandwidth restrictions. Extraction of AVAS data should be easier since it is designed for offload and archival. The quality of AVAS data will also be much better since it is designed to function without the GPS for periods of time when it does not have a good signal, and since it is designed to work in a virtual-signpost mode logging time of arrival at particular locations. This makes identification of the average and variance of run time between two stops in the system extremely simple.

Cross-check APC and AFC data [sched3]

Data from multiple collection systems can be used in a number of ways to audit data used in scheduling. APC data can be double-checked against AFC data for the same period, to assure APCs are not missing boardings, and that AFCs are not under-counting actual ridership. While an AFC failing to debit swiped cards is often an obvious failure which will be reported by a driver, other errors, such as a malfunction in an APC sensor will otherwise go undetected. Also, APC and AVAS or MDT data can be combined to audit the random sampling required to ensure APCs provide reliable measures of typical line performance. If scheduling begins to rely heavily on APCs for run times and loads, to continue to provide accurate schedules they should have some mechanism in place to assure sampling is representative. For example, a dispatcher could assign APC equipped buses systematically, rather than randomly. Or perhaps a particular operator or group of operators begin to deliberately avoid APC equipped vehicles. Either of these will deviate from the random sampling assumption, and both can be checked using archived MDT or AVAS data.

Estimation of OD matrix from APC data [plan1]

Soon the AVAS and APC systems will be on-line and the use of such data for providing route load profiles will be commonplace. Once all routes, or all routes in an area, have load profiles, this information can be integrated spatially over a number of areas to provide an automatic measure of the OD demand which is dynamic, changing daily as APCs are scheduled on new routes. This would allow the CTA to be extremely responsive to changes in demand. As data for load is directly calculated by the APC software, this would only require the integration in space of a number of different routes.

4.1.6 Application Selection

This step of the CTA Bus agency appraisal is a synthesis of the preceding information on expected agency benefits, expected complexity as a proxy for cost, mitigated by any expected synergies between multiple projects due to shared data products.

Application Description [ID]	Expected benefit	Expected complexity	Time frame
Emergency location system health monitor [sec1]	3	1	future
Line management usint BECS data [sm1]	5	4	now
Dispatch control using BSMS [sm2]	4	5	now
Trip-planner with reliability info [pi1]	4	5	now
Trip-planner with real-time information [pi2]	4	3	now
Real-time information for next-bus signs [pi3]	4	4	soon
Maintenance use of vehicle performance data [maint1]	2	3	future
Using system performance in marketing [mark1]	3	5	future
Identify precursors to unreliable service [anal1]	3	4	now
Load profiles based on APC data [anal2]	4	3	future
Performance auditing using BECS data [paud1]	3	5	now
BECS data for performance montioring [pmon1]	4	5	now
BECS / Demand data for performance auditing [paud2]	4	4	soon
Performance monitoring using supervisor PDA data [pmon2]	3	4	future
Archived BECS data in scheduling [sched1]	1	3	now
Archived AVAS data in scheduling [sched2]	4	2	future
Cross-check APC and AFC data [sched3]	3	3	future
Estimation of OD matrix from APC data [plan1]	2	4	future

Table 4.2: CTA Application Choice Matrix

To generate Table 4.2 I iterated through each agency need in Section 4.1.3, identifying individual projects under consideration. I ranked the agency benefit from 1 to

5 with 1 being some benefit to 5 being the greatest benefit. It is tempting to inflate rankings and give each project high marks, but as these values are to be used to select between projects competing for agency capital it is important to use the full range. In this process I took into consideration the agency's customer focus and ranked projects that provided a better level of service as more valuable than projects with benefits tied to labor-savings. Then while reviewing Section 4.1.5, I ranked project complexity from 1 to 5 with 1 being least difficult and 5 being the most difficult. This was directed by examination of the intermediate data products needed for each project, as described in Section 4.1.4. Lastly, the time frame for this project was identified, in this case amongst a small set of choices: now, soon and future. Projects labeled 'now' generally had data sources and hardware available. Projects labeled 'soon' had some data available but were lacking important storage, processing or distribution hardware. Projects marked 'future' require data collection hardware that has yet to be installed.

Application using BECS data Description [ID]	Expected benefit	Conditional Expected complexity	Time frame
Line management using BECS data [sm1]	5	2	now
Trip-planner with reliability info [pi1]	4	4	now
Trip-planner with real-time information [pi2]	4	2	now
Real-time information for next-bus signs [pi3]	4	3	soon
Maintenance use of vehicle performance data [maint1]	2	3	future
Identify precursors to unreliable service [anal1]	3	1	now
Performance auditing using BECS data [paud1]	3	3	now
BECS data for performance monitoring [pmon1]	4	3	now
BECS/Demand data for performance auditing [paud2]	4	4	soon
Performance monitoring using supervisor PDA data [pmon2]	3	3	future
Archived BECS data in scheduling [sched1]	1	2	now
Cross-check APC and AFC data [sched3]	3	3	future

Table 4.3: AVL-Available CTA Application Choice Matrix

In the process of reviewing intermediate data products, it became apparent that 12 out of the above 18 applications need position data for all vehicles on a route. Next those 12 applications' complexities are re-considered under the assumption that both real-time position and history of position of all vehicles on route, derived from BECS, were made available in an easy to access form to the agency. The results of this are

shown in Table 4.3. Given this data a range of projects become more attractive, with a high benefit to performance ratio; assuming other projects will follow on taking advantage of any new data products created, the position-using tasks' extra difficulty is justified.

In this case study, only applications that may be finished in a short time frame are considered. This eliminates all projects where the time frame is 'soon' or 'future'. From what remains, the greatest benefits are available from using Emergency Location derived AVL data for Service Management and Passenger Information. The use of archived AVL data for operations analysis also shows promise. These projects have the added benefit of simplifying future tasks using position data, including the Supervisor PDA project which is already underway. The results from implementation of these two projects are given in the following sections. While the actual agency costs were not certain when the application was selected, this prototype provides useful information on costs of distribution of real time data. This will be discussed in more detail at the end of the next section under the heading of Information on Requirements for Full Implementation.

4.2 Agency Benefits

To complete this agency case study, I implemented prototypes of the two applications identified in the previous section. I divide the following discussion of benefits into two sections – the first for a web-based real-time route-level vehicle position display tool, and the second for a post-processing analysis of archived data to aid in operations planning. As discussed in the Agency Appraisal, both of these tools use a common intermediate data product – position information for all vehicles along a single route – which is both archived and available in real-time.

In this system all the vehicles running a set of routes are polled using a proprietary interface to the Emergency Location system. The results from those polls are collected and entered into the agency standard Oracle database for redistribution and archiving. I identified a number of benefits of having position data as a frequently changing table

in the CTA standard Oracle database. The first was one of distribution. While bus position information has been available to Control Center users for some time via the vendor-provided interfaces for vehicle tracking it has been used infrequently and only in operations. By providing a standard database access to position history with near-real-time updates, other agency users became interested in such automatically collected data. CTA employees working in the planning department immediately began to access this position information from an Oracle client on their workstations, and were able to display the most recent position overlaid on a street grid using ArcView.

Another key benefit of database use is the ability to synthesize data that are logically connected, but not typically available together. A standard SQL query combines position data, operator data, and schedule data to provide vehicle position updates along with the run number and operator badge number, as well as using the schedule to determine the route [or possibly routes] of which that run number is a part.

4.2.1 Real-time Information Tool

This section presents the results from implementation of a real-time web-based interface to the position data described above. The route-level position reports derived from the AVL system are integrated into a dynamic display of vehicle position for the route(s) of interest. This could be used by supervisors to audit on-time performance as well as to monitor any deviations from routes taken by an operator. As agency employees are often also users of transit service, such a system is also an example of use of operations data for passenger information. Commercially purchased controller software has only facilitated tracking one vehicle at a time, while I created a system to illustrate the value of route-level AVL data. This system can be used in management of the line as well as for passenger information purposes. It also has the potential to more quickly identify vehicles with malfunctioning hardware.

Another category of user is the transit employee-as-customer. The route initially available on the internal website was the primary bus route used by employees at the

CTA's Racine staff office on their journey to work as well as for work-based trips. This prototype was designed in part to reduce unnecessary employee waiting at bus stops, but more broadly it was intended to build support for passenger information and related ITS initiatives.

In order to better explain this application, Figure 4-3 provides a snapshot of the interface. It is accessible from any web- browser on the CTA intranet, providing a visual display of recent and projected vehicle position for the route. There are two separate charts on this page, one for each direction of the route. The pluses on each zero axis are predicted positions of vehicles at the time of graph generation shown on the lower middle of each graph. This time should always be within thirty seconds of the current time, as the graphs are regenerated with that frequency on the server. The X-axis in the generated image is the position along the route, with street names every mile. The Y-axis represents minutes since the time of graph generation, in this case showing vehicle positions for the past 12 minutes.

The arrowhead indicates the most recent vehicle position report, and the dots connected by lines are previous actual position reports. The plus symbols on the zero Y-axis are the predicted "current" bus positions at the time shown in the graph, in this case 10:35:04. The solid lines connecting the most recent position report arrowhead and the predicted position plus symbol are presented in green on the website in contrast to the red lines of past vehicle trajectories. In this implementation this prediction was a simple projection from the rate of progress of the last two position updates. Prediction allows positions along the route to be compared at a glance, which is not possible using the unmodified position updates from the vehicles since from the nature of the polling, all vehicles do not provide their positions at the same time. The numbers that appear to the right of each arrowhead, and often overprint one another, indicate the vehicle number followed by the run number. A quick glance at such a graph shows, for example that a number of vehicles are "bunched" or running with very low headway, but it is not always easy to determine from the graphs exactly which vehicles are involved.

In order to more fully address this need, a text page is also available on the website

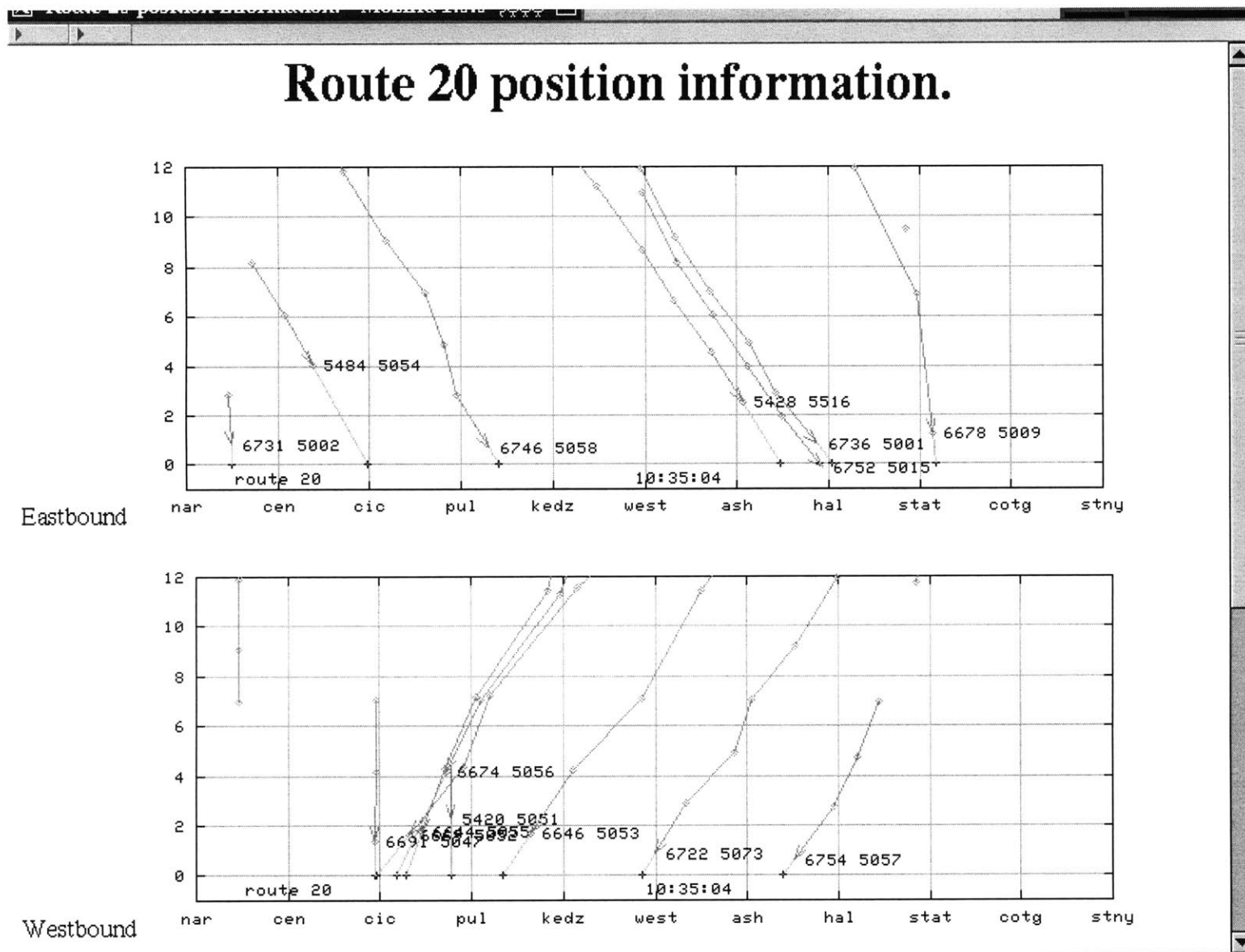


Figure 4-3: Position History Website

to provide a clear listing of information about each vehicle. This includes the vehicle's street address on Chicago's grid, as well as the vehicle and run numbers available on the graphical page. The direction of travel and vehicle speed are also listed.

10:37:29 Route 20						
20 East Loc		veh#	run#	dir		vel(mph)
5543 W	8 S	6731	5002	270	(E)	10.9
4855 W	7 S	5484	5054	270	(E)	5.1
3164 W	10 S	6746	5058	269	(E)	14.5
736 W	801 N	5428	5516	270	(E)	12.4
578 W	82 N	6736	5001	285	(E)	7.8
632 W	51 N	6752	5015	284	(E)	7.7
192 E	42 S	6678	5009	244	(E)	1.9
20 West Loc						
20 West Loc		veh#	run#	dir		vel(mph)
5182 W	1 S	6674	5056	89	(W)	11.0
4847 W	630 N	6691	5047	0	(N)	9.0
4961 W	6 S	6669	5052	90	(W)	7.7
4893 W	9 S	6644	5055	91	(W)	8.4
4172 W	667 N	5420	5051	101	(W)	0.0
4132 W	3 S	6646	5053	89	(W)	11.9
2957 W	8 S	6722	5073	90	(W)	12.7
1605 W	7 S	6754	5057	89	(W)	10.2

Figure 4-4: Position History Website

Passenger Information

The most popular use of the database information was a CTA intranet website which provided up-to-date graphs of vehicle position over time for CTA route 20, as shown in Figure 4-3. The author was a system user for one month in the summer of 2002. The CTA Racine facility is a 3 minute walk from a bus stop on route 20. Despite headways of 10 minutes or less in the late PM peak, waits of 20 minutes for a vehicle were not uncommon due to heavy vehicle bunching. Rather than wait, this author spent 30 minutes walking the 1.5 miles into the city, and on a few occasions was never passed by a bus.

An informal agency survey found 7 users not including myself who took advantage of the position information regularly. Of those using transit in their journey from work, everyone who was aware of the site referred to it before leaving the office. The fact that the site used simplistic projection algorithm did not seem to matter, the vehicle position information was the key. These knowledgeable and frequent transit users could make their own vehicle progress predictions which certainly would improve with experience.

Performance Monitoring

The processing tool can also run in an interactive mode which does not take the time to generate image files for all the possible plots that are used on the website. Instead it generates only what is specified on the command line. Updates are fast since they do not go through an image file intermediary, so the screen updates every second, showing a vehicle position history which smoothly scrolls backward in time. The most important feature of the command line invocation, rather than web-based use of this tool, is the ability to access arbitrary routes, rather than the small subset of routes available on the website. While limitations of the radio system allow only a few routes to be polled simultaneously, the database maintains a prioritized list of routes of interest. This table can be changed manually or through an easy to use interface; by changing the route table and specifying a route to the “posplothist” invocation it is possible to observe performance on an arbitrary route within five minutes, though routes with high percentages of non MDT-equipped vehicles will have inadequate coverage for performance auditing.

This tool is only a rough prototype, but such a simple tool permits some auditing beyond what is currently done for operators and maintenance workers. Within a thirty minute demonstration of this tool to a CTA bus operations employee, he identified and took down the badge of one operator passing a time point six minutes early, as well as observed two vehicles which were operating out of one garage although the computer records showed they were still officially based at a different garage. The vehicle which was extremely early was bunched with its leader, and also bunched with the leader’s leader which was several minutes late. This information is not being audited and confirmed as sufficiently reliable to write operators up for violations of rules observed through use of the position display system, but even in its current state it can direct the attention of those who are responsible for supervision. Mis-assigned garages show up on this system quite distinctly, as their run number has an uncharacteristic first digit, and the supervisors in charge of keeping records of bus transfers were notified that vehicles were not being properly transferred.

All these benefits were obtained without any hardware investment. The website was implemented using freely available tools and run on a surplus 110Mhz Sparc 5. The only costs were of this researcher's time in familiarizing myself with the agency data sets and programming, as well as the incremental cost of the use of a marginal amount of database space and server resources.

4.2.2 Post-Processing Data Analysis

Another use of automatically collected position data is the use of an archive of data to address questions of interest in operations planning. This application uses data collected from three weeks of operations, which provides a very comprehensive data set which covers each run multiple times. Using this data I identify factors that contribute to poor service reliability on a given route which better informs the agencies decisions on options to remedy developing service degradation before it is too late.

Trip Times

Figure 4-5 shows partial-route running times from three weeks worth of data for Route 77 with the solid line representing the hourly average trip time. This data set is unlike those collected by traditional manual run time observations. A well designed manual data collection program takes care to conduct a minimum of checks to answer basic scheduling questions: should we add or subtract a vehicle from this route to keep to our load standards, or should we adjust the scheduled run time to keep the schedule realistic. The automatically collected data provides sufficient data for weekend estimates, as well as for observing any change over time in route characteristics, and even assessing the daily predictable variations in run time which may have an impact on service quality.

The data from all weekdays are overlaid in Figure 4-6, which illustrates the worse performance of the PM peak, potentially due to exogenous conditions such as traffic. The large sample sizes available from automated systems for the first time provide agencies with the opportunity for statistically valid running and recovery time and

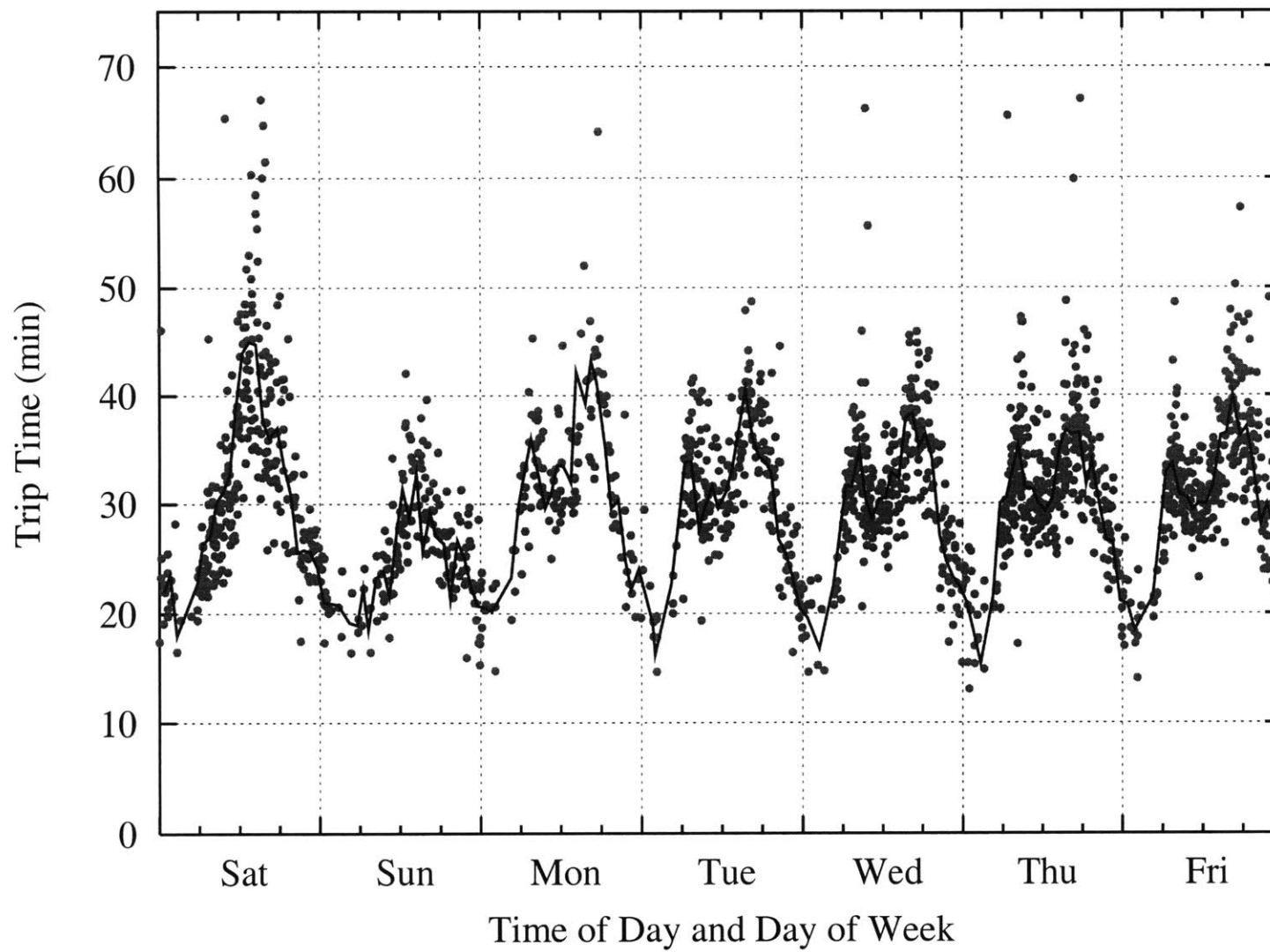


Figure 4-5: Trip Time on Route 77

schedule adherence analyses.

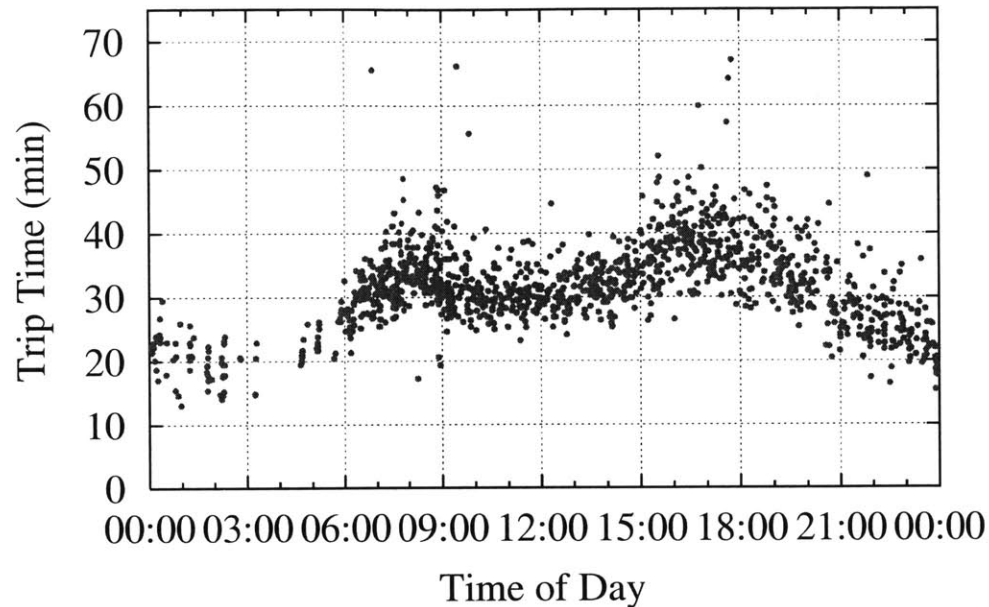


Figure 4-6: Weekday Trip Time on Route 77

On a segment of route 77 Westbound with a 30-35 minute run time, during a time period with 5.5 minute headways, 40% of operators are consistently more than 1 minute fast or slow, as shown in Figure 4-7. This shows that some operators are consistently fast while others are consistently slow. This is noteworthy in an agency with a policy against running early, although this analysis did not combine schedule information, which is not available in an easy-to-use electronic format. The AFC records for these trips exist, but were not combined to allow observations about typical vehicle loads due to the added complexity. Both these data source combinations would provide interesting avenues for extending this work.

Threshold Headway

A further observation using archived AVL data is that there is a threshold headway of seven minutes for significant service deterioration on this route on Saturdays. On the same route, the Saturday data shows that headways below seven minutes tend to degrade into very short headways, when comparing headway at an earlier location

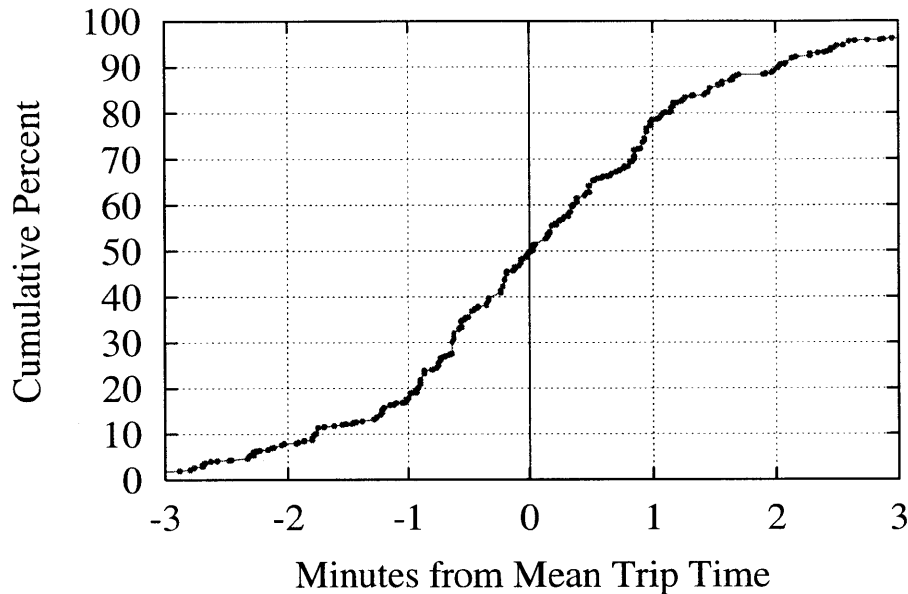


Figure 4-7: Operator Speed Variability on Westbound Route 77

[denoted dispatch] and a later location along the route [denoted terminal], as shown in Figure 4-8. The scheduled headway of 11 minutes is denoted by a green [or grey] line; most vehicles should have 11 minute headways: the point where the two green [or grey] lines cross. Headways do cluster around the intersection of the scheduled headways, however, if a vehicle leaves the terminal area with less than a 7 minute headway, it is likely to catch up to or even overtake the leading vehicle, generally arriving with less than a three minute headway. While it is impossible to separate the effects of any supervisory interventions, the Saturday service was chosen in part because the weekends have typically seen little service management intervention.

Data Quality

Archived data is helpful for determining the minimum data quality useful to an agency for a particular purpose. This section addresses sampling rate, data currentness and data reliability considerations. The normal Polled AVL data was augmented in this section by data gathered in a special-effort using the CTA MDT's. High resolution data was recorded on vehicles using floppy disks which were manually gathered after

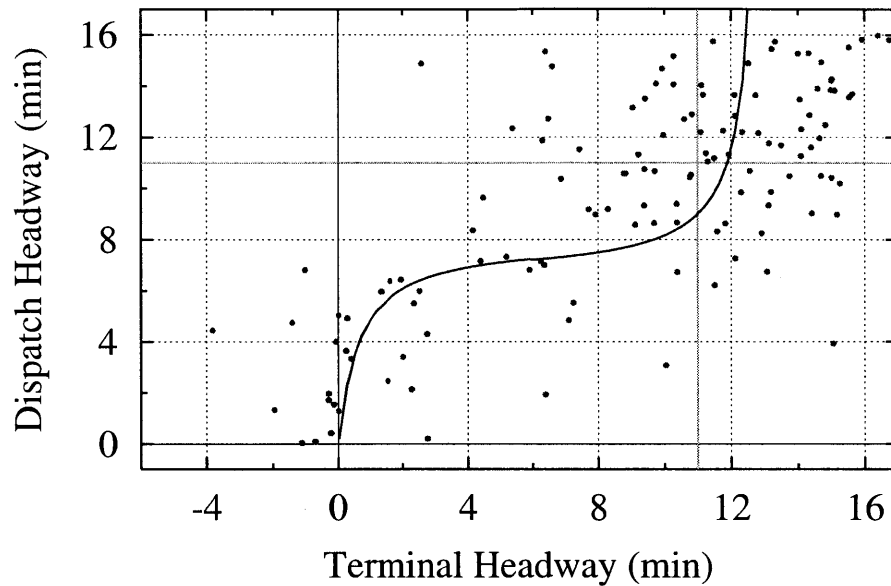


Figure 4-8: Saturday Headway Degradation

the data was written. While this data source is not one to rely on regularly, it can be useful to give a sense what questions better quality data available from the AVAS will help answer.

Transit data collection systems may log data on an event by event basis, for example an APC record when a user swipes a magnetic card or when a vehicle passes a signpost, or systems may simply log data at some pre-determined rate, such as an AVL system sending data over a polled radio data channel. In systems with timed data logging, important information can be lost if the data item of interest is not sampled often enough. In many desired uses, sampled data is projected to a known point, in order to facilitate comparison of information from one vehicle against another or of one vehicle against historical information, such as a schedule. Since some prediction or interpolation from observed values is required, the allowable sampling rate is dependent on quality of the underlying model used to predict and interpolate.

Further discussion will address the update rate of data items, the completeness of data sets and the accuracy of measurements, as they apply to typical agency uses of data. They will primarily address the effects as they apply to vehicle position data,

but the effects on the usability of passenger data is also discussed.

Update Rate Passenger data quality needs are to a large degree subjective and will vary based on expectations. The sampling rate is far less important to the user than other data collection system attributes, especially reliability. For pre-trip dynamic information, the user is likely to access the information several minutes before arriving at a stop or station. In this case, the predictability of the vehicle travel time dominates the reliability of the information provided. For wayside vehicle information, the user is in a position to notice any errors in predicted time of arrival, but when the vehicle is at some distance this too is likely to be dominated by the predictability of travel along a route. The AVL system's polling frequency, since it determines the average time since the last poll will keep the error in the predicted time in check, causing the error to decrease as the vehicle approaches. In order to identify the allowable poll frequency, what is acceptable to the user must be quantified, and a representative distribution of running times must be found.

Working with the following assumptions:

1. arrival time forecasts with an effective accuracy within 50% or $+/- 1.5$ minute – whichever is greater – will be acceptable.
2. the default coefficient of variation of 0.16 for sub 20 minute run time trips from the Transit Data Collection Design Manual, [40] as a proxy for the short-segment run time variability expected as a vehicle approaches a stop.
3. run times are normally distributed and that the actual accuracy of the base data item $[+/-50m, \text{ for example}]$ will be so low as to not influence the estimated value.

It is determined by the poll cycle where the effective accuracy available by the estimate $[\text{in this case } 2 * \text{remainingtime} * 0.16]$ is just marginally greater than the minimum value $[1.5\text{min}]$ which is in this case the poll made when the vehicle is approximately 4.7 minutes away. The amount of time it takes the percent error to become just marginally below its maximum allowable error $[50\%]$ determines the poll rate for the given error levels and operating characteristics. In this case the minimum poll cycle

is 1.7 minutes or about 100 seconds. On routes with more variability – a higher coefficient of variation – the polling rate would have to be higher in order to provide the same reliability of position prediction.

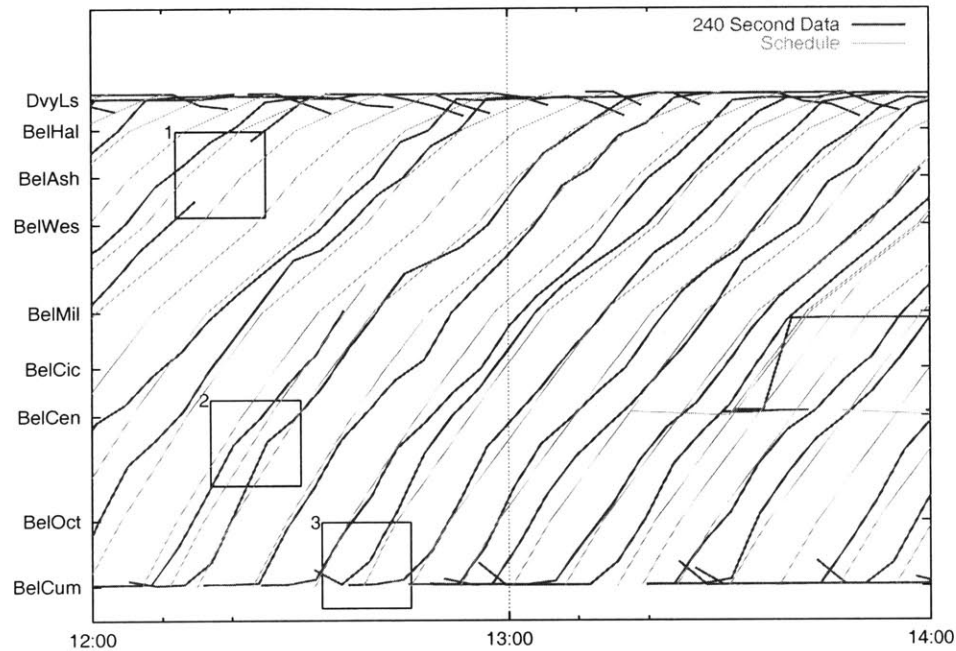


Figure 4-9: Example Route 77 Eastbound Bus Data / 4 Minutes Polling Cycle

As shown in Figure 4-9, data polled every four minutes, can be useful for line management and performance auditing. Bunching is clearly evident, with the close parallel tracks indicating two minute headways inbound at the time-point labelled BelCic at 12:50. Running ahead of schedule, contrary to CTA policy, is also evident with a number of vehicles departing the terminal early, for example the one shown in Box 3. However this data is insufficient for investigating system characteristics, such as identification of the contribution of dwell time to run time. For example in Box 2, the origins of the dramatic change in speed just before the point BelCen are unknown. Other characteristics, such as position reporting reliability, are important if data with such low polling frequency is to be used for management. One problem, illustrated in Box 1, is that a single invalid [out of bounds] data point can lead to a large gap in position understanding. After the fact, this can be approximated with a straight line, however if this data were being used in real time, the position would

have to be predicted for eight minutes, which represents more than one mile of vehicle travel.

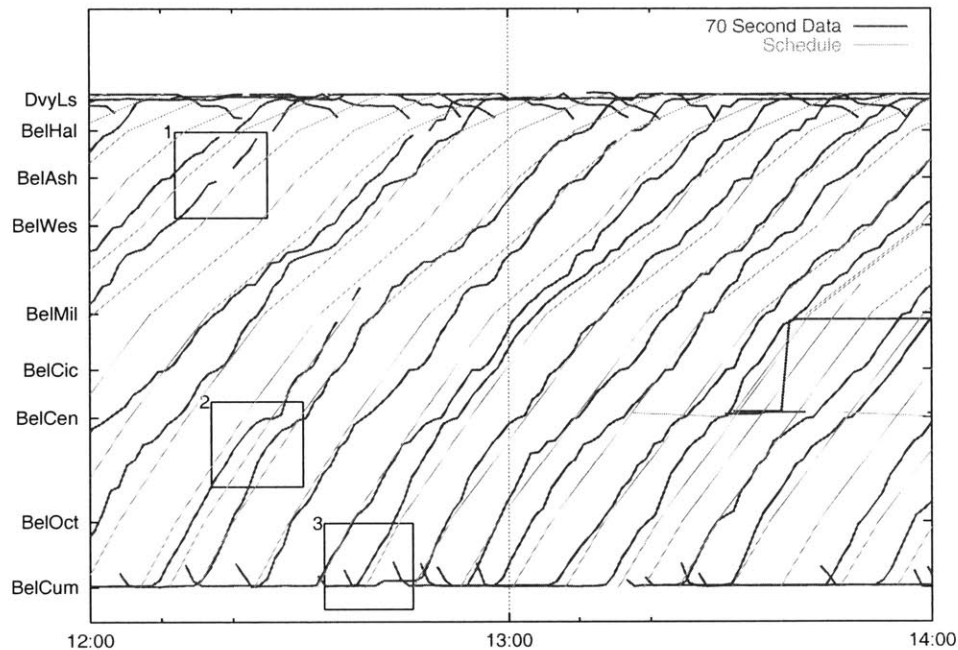


Figure 4-10: Example Route 77 Eastbound Bus Data / 70 Second Polling Cycle

Data reported approximately every minute as shown in Figure 4-10, give substantially more detail about operations. The leading vehicle shown in Box 2 is delayed for at least five minutes at BelCen as are a number of other vehicles at other times in the afternoon. In Box 3 it is apparent that the vehicle did not leave as early as it appeared in Figure 4-9; in fact, the 13:00 scheduled departure was substantially earlier, yet looked equally early from the previous data set. Lastly, the missing data in Box 1 is much less dramatic and prediction of vehicle progress through this section is more practical, though progress past BelAsh is especially variable in this instance.

High resolution data obtained using on-vehicle recording capabilities shown in Figure 4-11 illustrate the benefits of such collection for analysis. Positions are available at a high enough frequency to identify individual stops. The Detail section outlined in the center is shown in Figure 4-12. Such stop-level high resolution data is representative of the quality of data required for traffic signal priority. The arrival time prediction and intersection exit detection must be precise within seconds in or-

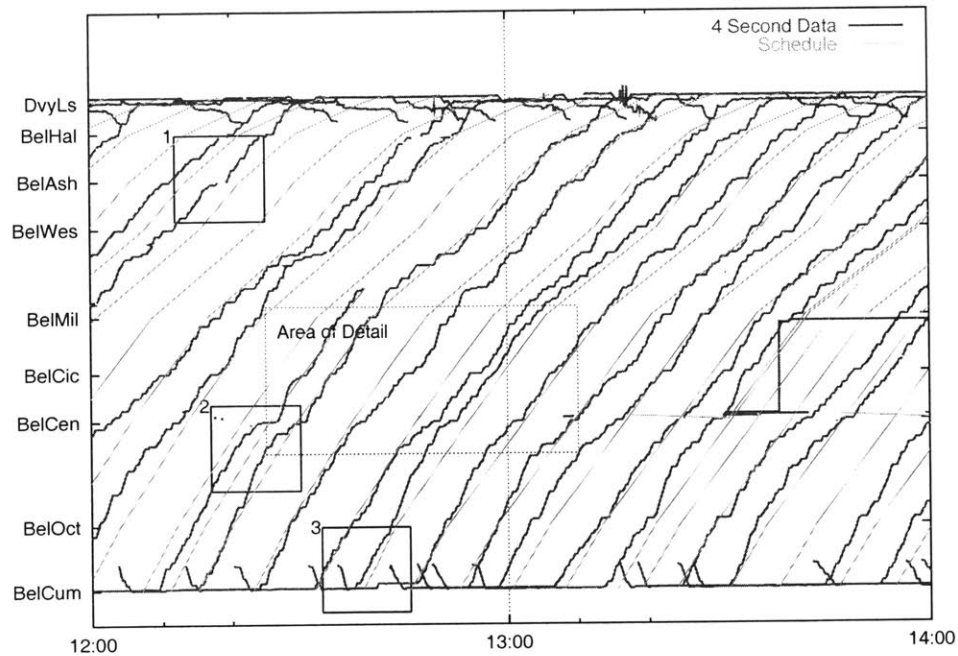


Figure 4-11: Example Route 77 Eastbound Bus Data / 4 Second Polling Cycle

der to allow the bus through but minimize disruption to automobile traffic. Such high resolution data required in real-time necessitates that vehicles communicate directly with equipment in the intersection, rather than a centralized solution requiring communications to the central repository and back out to individual street lights.

In this detail section, the recorded door open/closed data is overlayed as plus symbols for each record with a door open. Such recorded data provide agency analysts with the opportunity to separate dwell time from traffic delays. Such information can be useful in making the case for a dedicated bus lane or for particular Bus Rapid Transit alternatives, such as pre-paid boarding at a particular stop. [4]

Completeness For each of these levels of detail, the data set suffers because there are a number of missed trips in this graph. This makes this data ill suited for headway variance analysis, which is a serious deficiency in the analysis of a high frequency bus line, where many users will not time arrivals to the vehicle schedule, but will instead arrive randomly in time expecting only a short wait. These data sets with gaps are still useful, however, because they do allow observation of vehicle bunching in systems with high headway variability. While a lack of observed bunched vehicles

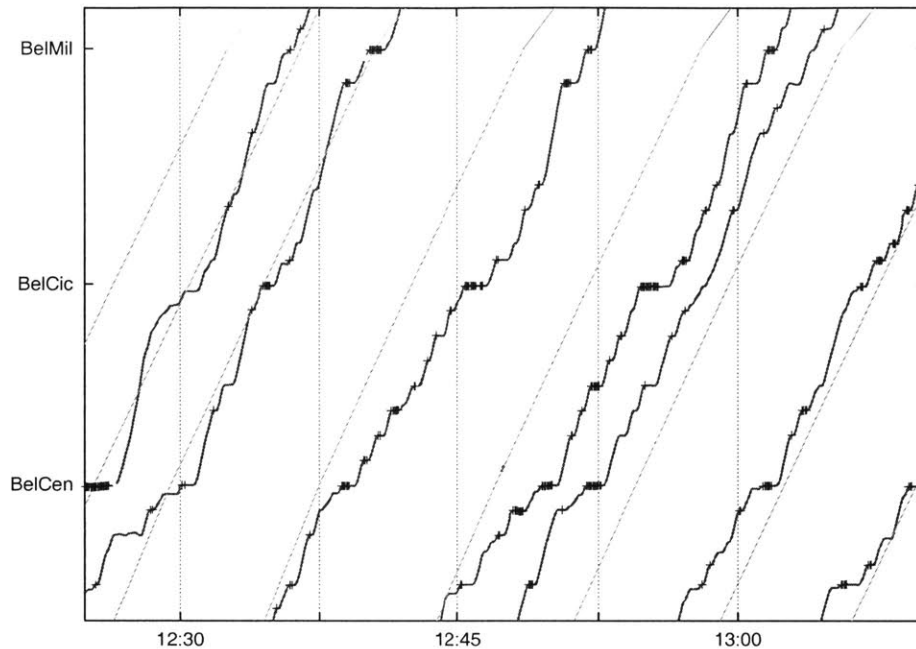


Figure 4-12: Detail from Figure 4-11 with Door Open Tick Marks

does not imply high quality transit service, the existence of bunches argues that service reliability might be improved.

Data completeness on another time scale is of concern to the deployment of APCs in the CTA. Given the CTA procurement of APCs for 14% of the fleet, and less than 100% of the fleet will be operational at any given time, the number of routes sampled will be adequate for statistically valid estimates of passenger figures, but each trip on each route will only be observed an average of once per month, if APC vehicles are distributed without route, time or operator biases. This is not enough to give detailed operations data on individual movements. It is for this reason that the AVAS system was designed also to log high quality position data.

Accuracy Data used to ensure operator security, such as the current position of a bus, ideally should have sufficient precision to direct an emergency vehicle to within visual range of the vehicle. It should be able to achieve this level of accuracy with high confidence and high availability. This requirement is not particularly stringent; emergency vehicles can accommodate at least a city block of ambiguity – 200m for something so large as a city bus in distress. However it should be able to achieve this

level of accuracy with a high confidence – 99% – and with high availability – 99% – as the consequences of failure are severe. Such silent alarm position reports are not likely to be needed frequently, but when they are needed the time between the request for position and report of position should be minimal – under one minute.

A plausible basis for such recommendations is that of the E911 cell-phone location requirements.[38] Unfortunately these standards appear to be influenced strongly by technology. The rules state that 95% of new phones must be enabled by October 1, 2002 and cell phone carriers must achieve 50-m accuracy for 67% of mobile emergency calls and 150-m accuracy for 95% of all calls. The required accuracy distances are doubled for network-based location technologies where the cell towers triangulate a signal rather than the handset determining its own position. No availability figure is specified in the FCC cellular telephone requirements.

For service management and performance auditing, vehicle position requirements are similar. The position of several vehicles along a route is needed at higher frequency [see the discussion above on update rate] at multiple points en route, with sufficient accuracy to observe headways to 1 minute accuracy – which translates to approximately 100m given uncorrelated errors and 15km/h average speed. Confidence need not be as high as in an emergency situation and a lower value such as – 90% – can strike a balance between data usability and system cost.

Information on Requirements for Full Implementation

These applications provide a great deal of information about communications, processing and storage requirements for a future CTA vehicle position collection, distribution and archival tool. This information can be used to help improve cost estimates.

It is important to understand the details of the CTA bus emergency location system in order to illustrate the data flow into and out of this new system component. A machine known as the Communications Processor [CP] in the Control Center is central in all bus data communications. The CP communicates with a bank of digital radios, using three antennas installed around the city to provide full coverage in the service area. The CP polls each vehicle assigned to a channel once every 40 to 70

seconds, depending on the number of vehicles on a given channel. When polled, a vehicle can respond that it has an event to report, which triggers a communication from the vehicle, for example an operator log-in event would result in a message including run number, badge number, time and vehicle position. In the current system, if a vehicle has no event pending, the vehicle location is not returned, however an enhancement ordered in Summer of 2002 will provide a vehicle position report on every poll. When implemented this will switch the CTA's system from an exception based system to a true AVL system. Messages can also be sent from the control center to individual vehicles. Such messages include text messages to be displayed on the MDT to provide information to the operator, as well as requests for the vehicle position on the next poll. [28]

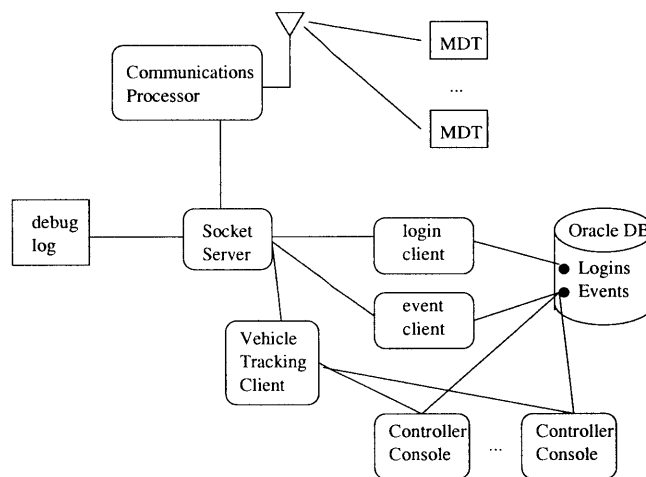


Figure 4-13: Bus Communications Data Flow Diagram

Vehicle messages pass from the CP to a machine known as the Socket Server for distribution. A graphical representation of the connections of the different components is shown in Figure 4-13. Clients use this intermediary to send messages to vehicles and request particular types of return messages. One such client is a vehicle tracking tool which could potentially be used in a vehicle hijacking situation. This tracking tool puts that vehicle in a high-rate-polling state, where it gets priority in the poll cycle and is allowed to return position reports every 15 seconds. The most recent vehicle position and time report is kept internally on the Socket Server and

made accessible through graphical and command-line interfaces, but not in a general purpose database. Another Socket Server client is an interface to keep a database of operator login information. This vehicle login database keeps a record of the operator and run number logged in to each vehicle, with a corresponding day and time. Messages which require controller attention generate events; events are entered into an event database and are archived for a short period to facilitate incident investigation. [2] The events and operator login data just described are presently kept in tables in a general purpose Oracle database, and efforts are underway to adopt this Oracle database as a central repository for operations data.

The development server used in this project is on the same local area network as the Socket Server. Communications with the Socket Server are relatively easy compared to the communications from the Socket Server to the individual vehicles, which must pass over this time-divided data link. The radio data link will be the most difficult to improve constraining factor in the input of data for this case. The vehicle position data is retrieved from a local position store on the Socket Server, and added to a database running on the development server.

In this case the two dynamic data inputs added to the database are the current list of vehicle run numbers and the reported vehicle positions. Vehicle position reports are fairly compact, fundamentally just two floating point numbers as well as a time and a vehicle number. The position report message size is 32 bytes. Even if all 1600 MDT equipped vehicles reported their position every 60 seconds, this would require approximately 30 position updates per second, or a data transfer rate of 900 bytes per second. The run number list is updated as operators log in and log out, entering new values into the MDT. Given a pessimistic average run duration of 2 hours, over the course of a day there will be fewer than $12 * 1600$ or 20,000 updates, for a rate of no more than one changed entry every four seconds. These records are already available in the database, so no database inserts are needed, only queries. On such a small table, the current hardware can handle queries every few seconds without appreciable overhead.

The outbound data flow from the database is of a substantially different charac-

ter. The most demanding user is the bus controller, who needs several minutes of position history for all the vehicles on a route. Fortunately this is only around 30 vehicles, and 15 minutes of history is approximately 15 position reports per vehicle. With 10 second updates this is only 90 vehicle positions per second. To accommodate 10 controllers simultaneously accessing different routes would require 900 vehicle position records per second. This is a substantial demand which could not be accommodated on the existing hardware. Another user is the web passenger information user. In order both to allow a large number of users and to maintain system security, the passenger information should be served by a second machine located outside the agency firewall. This machine needs recent vehicle positions pushed to it as often as possible. This is at most a record for each vehicle in the system [1600], every 10 seconds, or only 160 records per second. The server used in the current prototype, a Sun E-450, was operating at approximately one third of its capacity when integrating and distributing position reports from 70 vehicles, so it should be able to accommodate approximately 1/10 of the full vehicle load with this mechanism for indexing, archiving and distributing vehicle position. More recent, high performance computing hardware should be able to handle this load without difficulty.

4.2.3 Summary

On the whole, the trial use of this system illustrates a number of benefits available to the agency. These observations of vehicle performance, and future observations enabled by the archival of bus position data, can lead to a number of operational improvements. Improved schedules, improved enforcement of policies against departing time-points early, and details about the sensitivity of routes to natural perturbations, which may lead to revised operating policies and or a push for more effective service monitoring and management.

The polling of a set of vehicles to get information for several vehicles along a route can be an effective means of more fully utilizing an exception-based AVL system. In the case of the CTA, where the agency is in the process of transitioning to a full AVL system, this serves as an example of some of the applications desired in such a

system, as well as some of the pitfalls to avoid.

4.3 Implementation Barriers and Solutions

This section presents the solutions to particular barriers to implementation encountered for both applications. The barriers encountered include several of those discussed in Section 2.7. Specifically, I encountered a number of data formatting conventions that were inconsistent, undocumented and of restricted distribution. To deal with these problems I reformatted the inconsistent, undocumented and proprietary formats into a common format stored in the database. I also had to initiate data archival for a number of data items of interest in order to overcome the lack of storage of passed data sets. I overcame a lack of computer processing power and communications bandwidth by focusing on a small subset of agency routes, and taking advantage of Chicago's grid-based street lay-out. Lastly, I overcame a pervasive skepticism about the usefulness of installed Emergency Location system by focusing on what it can provide rather than dwelling on what it cannot.

4.3.1 Define Consistent Interfaces

This section addresses the solutions which facilitated extraction of useful information from the sources of data formatted in inconsistent, unpublished and restricted-distribution ways which were all encountered in the course of application development. Key steps in this are to share the data by providing it in an agency-standard Oracle database for desktop use by anyone interested. It is important to consider all potential agency users of system data. The primary users of data from these initial applications are the passengers, bus controllers and Bus Operations Planners. As was shown in Figure 4-2 there may be some use of expanded bus data collection in Scheduling and Performance Auditing, and potentially even to augment the current Security functionality. To aid in future efforts in planning, the position, run number and operator id should be archived. For future Performance Monitoring purposes the data should be available in near real time in an easily accessible form. To aid

in Security the data should be available in real time, in a manner amenable to reliably identifying off route vehicles, as well as for use as a security hardware reliability auditing tool.

Some issues encountered in data use fit into the category of unfortunate conventions. One such convention is that operators do not presently enter a route number into any system that is automatically transmitted in real-time. Another is that run numbers are not unique across garages; and the system-appended garage identifier is sometimes incorrect when buses switch garages and the system isn't kept up-to-date. Furthermore there is a convention of the use of letters as single character designators for agency garages; some systems do not handle letters and instead substitute numbers, using a mapping of the names of CTA Bus garages from letters to numbers. This section addresses methods for accommodating these conventions and suggests ways to improve each in the long term to avoid unnecessary complexity in automated processing.

Another issue is that a number of systems encountered were not designed to provide a convenient interface for interaction with external programs. Some of the data used in the system are not available in machine-usable electronic form and need to be pre-processed. Restricted access to data format specifications were encountered with both the scheduling software database as well as the communications to and from BECS, requiring a Non-Disclosure Agreement [NDA] with the vendor, Orbital Sciences, to fully use the system. Even the CTA Oracle server presented difficulties with interactions with external programs. For each of these I provided a layer of abstraction above the raw system level communications in order to facilitate other uses.

Accommodating conventions

The process of determining the correct vehicles to poll was not in practice as simple as this description suggests. Difficulties were encountered in working with the information available at the CTA, and are likely to be encountered elsewhere. The first minor difficulty is that the schedules department uses single character garage

identifiers, specifically: 1, 6, 7, A, 5, F, K, or P. In operations, a single digit from 1 to 8 is used as a garage identifier. The mapping of garage identifier from schedules to operations is known by people in operations, but is not widely known in the schedules department. Written documentation on the mapping is difficult to find.

A further difficulty is that the operator enters only a three digit run number, but this is not unique without reference to the garage. Electronically, the current garage for a given vehicle is stored in the database and is kept up to date by garage personnel filling in an electronic vehicle transfer form. The bus's current garage is added to the manually entered run number, however that may not be the actual garage from which the vehicle was dispatched that morning. Vehicles also have a home or assigned garage, which is where they are supposed to be kept overnight. As a mechanism for identifying these mis-identified vehicles, (caused by current garage records not being kept up to date) those vehicles with run number matches using the assigned garage rather than the current garage are also polled. The query actually used, which combines my responses to each of these issues is given in Appendix A.

Abstraction

The position reports returned from vehicle polls are logged by the BECS system to a text file in a proprietary message format as received over the air. This log file contains all system message traffic, and had previously only been used in the course of system debugging. Collection of polling data began with a system to process the log file nightly, selecting the location messages and extracting position, speed, direction and time in a format that is easily machine-usable. The list of vehicles polled and operator login information – badge and run number are collected separately. The use of this archived data is discussed in Section 4.2.2. This separate collection and archival was the source of some difficulty, as some data sets are missing the historical record of run number and operator information, making the vehicle positions much less useful.

Over the course of the creation of this tool, the vendor of the AVL system delivered a new utility to access the most recent position report from each vehicle. Unfortunately, the interface provided to this internal representation is limited in its

flexibility as it provides no position history or route-to-vehicle mapping. In order to provide bus position information as an agency-wide resource, I began automatically gathering position data from the proprietary system automatically entering updates into a table [veh_position_history] in the Oracle database every ten seconds. This system keeps time stamps of database entry as well as the vehicle's time stamp for the actual time of data collection. Typical weekday operations generate approximately 80,000 position reports. Due to database size restrictions, the database is pruned every night to retain only the previous three days' vehicle position reports, however the database is dumped to flat files nightly for archival purposes.

A larger difficulty is that the schedule information containing the translation from run number to route number is part of a proprietary scheduling package in a separate database with a proprietary / undocumented schema. Initially, the schedule was only available in a text format geared for legibility to humans, but not for ease of use in further processing, though employees are in the process of improving access to this schedule information.

Another factor impacting data usability is the mapping of schedule and vehicle login to a particular set of vehicles to poll. The route number to vehicle number mapping is not available directly. However, the run number to vehicle number mapping is already maintained by BECS in an operator logon database table. The schedule information containing the translation from route number to a list of run numbers is part of a proprietary scheduling package; by manipulations of the human-readable schedule output, rather than an interface to the scheduling package, a mapping from a route number to all the run numbers possible on that route was also placed in the database. This has the difficulty of requiring intervention to assure synchronization with the current schedule. For a given route, a simple query provides the vehicle numbers of vehicles where an operator has indicated the run number is on one of the list of routes of interest.

A final detail on database access relates to the fact that there are actually two parallel operator logon tables, one on each of two database servers. In the rare event when one machine fails the other identical machine is available to replace it. Unfor-

Unfortunately, the system was not designed to interact seamlessly with outside programs. In order to continue polling the correct vehicles and collecting responses, I needed to keep track of which server is currently active. I created the tables as pointers – synonyms, in the Oracle vernacular – to the true login and vehicle tables. For continued operation despite server fail-over, only the synonyms need be changed on the polling system to point to the tables on the working server.

4.3.2 Working With Available Resources

The computer power and communications bandwidth were important constraints in the polling of data and its real-time display. As the communications channel can only accommodate 90 vehicle polls per minute, I select ones to poll using a route priority list. In order to save computation time the system mapping to street coordinates was optimized to take advantage of Chicago's grid street system. All systems used spare processing power on existing computers and outdated agency surplus computers. The project time limitations led to the use of commonly available flexible software tools to do many of the hardest tasks.

Prioritize use of available resources

In this case study, the two most constraining factors were over-the-air radio channel bandwidth, and computer hardware resources for the real-time operations portion of the application. The computer hardware constraints were not binding for the post-processing portion of the application since the calculations need not be done interactively. Polling position of a larger number of vehicles leads to a longer poll cycle than for a smaller number of vehicles, since the radio channel is used for some period of time for the transmission of coordinates. CTA management made the estimate that polling 100 or fewer vehicles would not significantly lengthen the poll cycle. In order to fully utilize the bandwidth available, multiple routes of interest may be specified for polling. However this can potentially lead to a set of vehicles of interest larger than can be polled over the radio data channel. In order to accommodate

that, each route is given a priority, and the list of vehicles is sorted by the priority of the route on which the vehicle is traveling. To restrict the impact of this system on normal operations, only the 90 vehicles with highest route priority are selected for polling. The actual polling is accomplished using a pair of UNIX Korn shell scripts which selects the 90 vehicles of interest, then properly formats a broadcast vehicle command message with a request for position to those vehicles. The vehicles are polled in this manner every 60 seconds.

Take advantage of area specific information

The constraint on computing hardware for the real-time portion of the case was to share the development hardware on hand, specifically a two processor Sun Enterprise 450 server with 512 MB ram and 60 GB disk space. The available server is typically lightly loaded by its other duties. A potentially very difficult step in processing each record is mapping coordinates to a particular route. One candidate mechanism for mapping to position along a route is to map each point to the closest route segment, and determine distance along each segment traveled. In the interest of processing and implementation time, a mechanism which takes advantage of Chicago's grid street system was used. Acceptable results are possible for routes that are primarily East/West, primarily North/South, and even smoothly diagonal routes by utilizing just the Latitude or just the Longitude coordinate. To further speed implementation, approximate conversion factors were used to translate from Latitude and Longitude into a Chicago-centric North and East coordinate system with 800 units equaling one mile. This largely agrees with the city block numbers, with a manageable list of exceptions. This does not provide reasonable street addresses for diagonal streets, but vehicle progress on buses routed on diagonal streets, can be adequately observed by focusing on either the Longitude or Latitude component of bus position.

Use surplus computers

A range of different computers were used in development of these projects. The most computationally intensive task is the insertion of position records into the Oracle

database as a clearinghouse for other data users. In order for the table to be useful for line visualization it should hold more than the most recent position update: it should also have some limited history easily accessible. An insert can be a very quick database operation, but in order for new records to be usable the table must have indices which are kept up to date with each insert. When the database of recent inserts grew beyond a few thousand records – the equivalent of only five minutes worth of data if all agency buses were providing data – system queries which initially took less than a second on average began to require over one hundred seconds of processing. The careful selection of database keys and the creation of database indices [see Appendix A for specifics] reduced processing time to approximately two seconds from a database with over 100,000 records. In the database community it is well known that proper indexing is always good practice and crucial for performance in large tables, but I encountered difficulty finding people with this level of system knowledge, and had to do my own research into database tuning.

Use standard, flexible data processing tools

For both real-time and post-processing tools, position information is dumped from the database to flat files which are processed by a script written in perl, a flexible data extraction and text processing language. The perl script for generating web-pages sends commands to a freeware program called gnuplot to generate images copied into the web hierarchy of a stock apache web server installation. The perl processing tool “posplohist” maps from the GPS derived Latitude and Longitude into a position along route index for plotting, as well as calculating the speed and direction of progress, all simple mathematical operations in this implementation. Perl is very good for simple operations done line by line on large files. The preprocessing also makes an estimate of the route variant that a particular vehicle is operating, currently limited to a Northbound / Southbound or Eastbound / Westbound split.

4.3.3 Working Within Collection System Limitations

The Emergency Location system has had a number of setbacks in its installation at the CTA. The radio channels used for data reduced the bandwidth available for voice communications, reducing the effectiveness of the primary means of information exchange for the operator. It has also had reliability problems; MDTs often fail in the field. This has led to a general dislike of the system and skepticism of its usefulness.

The applications developed in this case were successful because they avoided system weaknesses, and they were not biased by unrealistic expectations of the Emergency Location system capabilities. The MDT performance is notoriously bad downtown due to the urban canyon effect. As downtown is the terminal for a number of routes, the system is not appropriate for full-route run times and analysis of terminal dispatch and lay-over practices. However as the post-processing application shows, by using points along a route near terminals but before data quality drops off, a large quantity of high quality segment-level running time is available for analysis.

The importance of approaching data sources with an open mind can not be stressed enough. While it is reasonable to be wary of systems with reliability problems, there is often useful information available from data sets with incomplete coverage. Summer interns, such as this author, are good candidates for exploratory and experimental projects with relatively high risk, but with the promise of great rewards. There is nothing like a useful application such as the position website to illustrate that a system does have some promise as a source of agency data, even if not exactly as originally intended.

Chapter 5

MBTA Application

This case describes the application of the framework in Chapter 3 to the Massachusetts Bay Transit Authority Rail division. This chapter begins with an appraisal of the agency's ability to meet its functional needs, and identifies a promising software application to improve it. This application combines heavy rail train position information and historical demand data to provide management with daily line performance summaries from the passengers' perspective. The next section of this chapter describes the improved agency performance using this mechanism for supervisors to audit system performance and gauge the effectiveness of interventions designed to restore service after a disruption. It closes with a discussion of decisions made in implementation that may be helpful in future projects.

The MBTA is the fifth largest public transportation system in the United States, with over 700,000 rail boardings per day. The agency operates more than 635 light and heavy rail cars over its 185 total track miles. (See Figure 5-1.)

5.1 Agency Appraisal

This section first reviews the data collection hardware currently in use in the agency, as well as hardware in procurement. Next the available or soon-to-be-available hardware is compared with the current uses of data sources in a graphical format to help identify especially promising areas for helping the agency meet its functional needs.

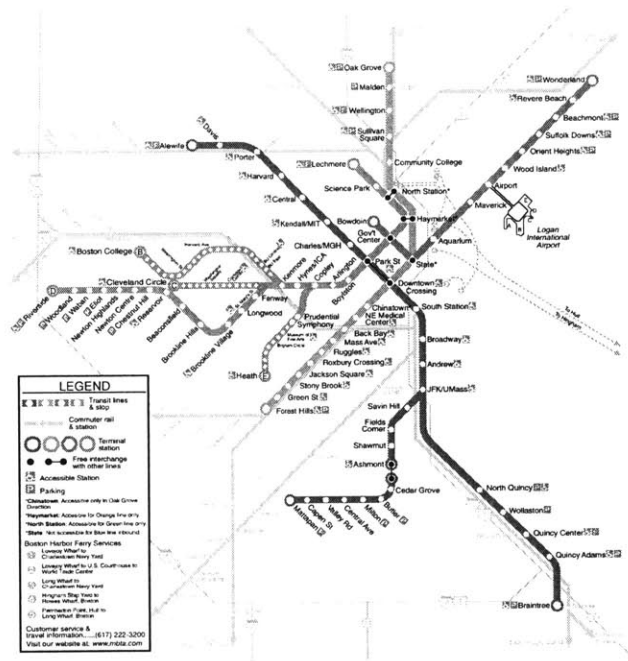


Figure 5-1: MBTA System Map

Then comes an appraisal of the agency's ability to meet a range of needs, focused on those needs with potential for improvement through increased use of automatically collected data, from the short-term service management decisions in operations to the longer-term planning decisions. Specific options for improving agency fulfillment of each of the listed needs using automatically collected data are identified. Intermediate data products used by multiple applications are identified and the difficulty of information extraction is addressed. Finally the challenges presented in the combination, distribution, and display of these data products are presented. Based on the benefits and complexities in light of knowledge of agency values and discussions with management, one application with potential to benefit the agency emerges.

5.1.1 Agency Data Sources

The MBTA has data systems which provide a great deal of information on train position, but does not yet have any automatically collected measures of ridership. Such data will be available in the future as the MBTA has awarded a contract for an Automated Fare Collection system with a June 2005 scheduled completion date.

The data available on train position differs depending on the line. For the three heavy rail – Red, Orange, and Blue – lines, high quality data is available from the Automated Train Control system. For the light rail, Green Line Automatic Vehicle Identification data is available from thirty-one detectors to provide automatically collected position reports. No automatically collected position data is available for the light rail Mattapan trolley.

The heavy rail ATC system’s primary responsibility is to ensure safe spacing between trains to reduce the possibility of a collision. It tracks trains by the combination of AVI-type detectors along the route and track segment level occupancy information from wayside bungalows. Strategically located AVI detectors at particular positions along the lines provide a comprehensive list of all the cars running together in a train consist. Train tracking software uses the track segment occupancy information to project these consists forward as the train progresses over a sequence of track segments. The ATC system also provides the information for real time train position displays in the MBTA’s Operations Control Center. Train position information is logged into a database, and kept for a short period of time to aid in incident investigation.

The Green Line light rail AVI system was installed for automatic train routing among the line’s four branches – B, C, D and E. AVI data is relayed back to the MBTA Control Center for management and auditing, but the responsibility for service spacing and regularity lies with operators and supervisors at key locations along the right-of-way. The coverage provided by the AVI detectors does not give a comprehensive picture of the entire light rail system. For example, the last outbound detector on the E branch is $1\frac{3}{4}$ miles from the end of the line, between the Symphony and Museum of Fine Arts stops, and the first inbound detector $\frac{3}{4}$ miles from the end, near Brigham Circle. This makes AVI data less attractive for further data analysis than the ATC system on the Red, Orange and Blue heavy rail lines.

The quantity and quality of data available are much greater than in the CTA bus case. This is due in large part to the fixed nature of the infrastructure on rights of way owned by the agency. The MBTA rail ATC system has direct serial communi-

cations with each of the computer controlled power distribution systems in wayside bungalows. Each time a track segment becomes occupied or vacant, a track on or off message is sent to the control center. The systems in bungalows also digitize auxiliary information primarily for control center display including departure bell ring time, exhaust fan activation state and switch state. Track segment lengths range from as short as 50 feet in some crossover areas to as long as 2000 feet in high speed, uncongested, sections. The on/off state of each track is instrumented by the electronic wayside power controller and transmitted in a serial data stream to the control center for monitoring. For safety reasons the system assures that at least one empty block must be maintained between successive trains. Thanks to this lack of position ambiguity, the MBTA train tracking system is able to reliably keep track of the identity of all vehicles in the system.

The most reliable measures of demand presently available are from labor-intensive and infrequent passenger counts conducted by the Central Transportation Planning Staff [CTPS]. As specified in the February 2003 contract with the German firm Scheidt and Bachmann, this will, however, soon begin to change. Work is scheduled to begin on installation of an Automated Fare Collection system for Rail and Bus replacing the 475 turnstiles at 90 subway fare collection locations in June 2004, with completion only a year later. This will lead to the phase-out of tokens and the phase-in of magnetic as well as smart-card fare payment. The contract also includes infrastructure for connection of the newly installed turnstiles to a centralized computer and software for providing fare statistics.

Another very different source of demand information is the MBTA trip-planning website. The site allows you to get information for stops and route schedules near a given address or landmark as well as traveling directions between two such locations. While not everyone who submits an inquiry actually intends to use transit near a particular location, or even intends to travel to or from that location, the information on web-site users' requests is a source of automatically collected data which may be useful.

5.1.2 Data Source Utilization

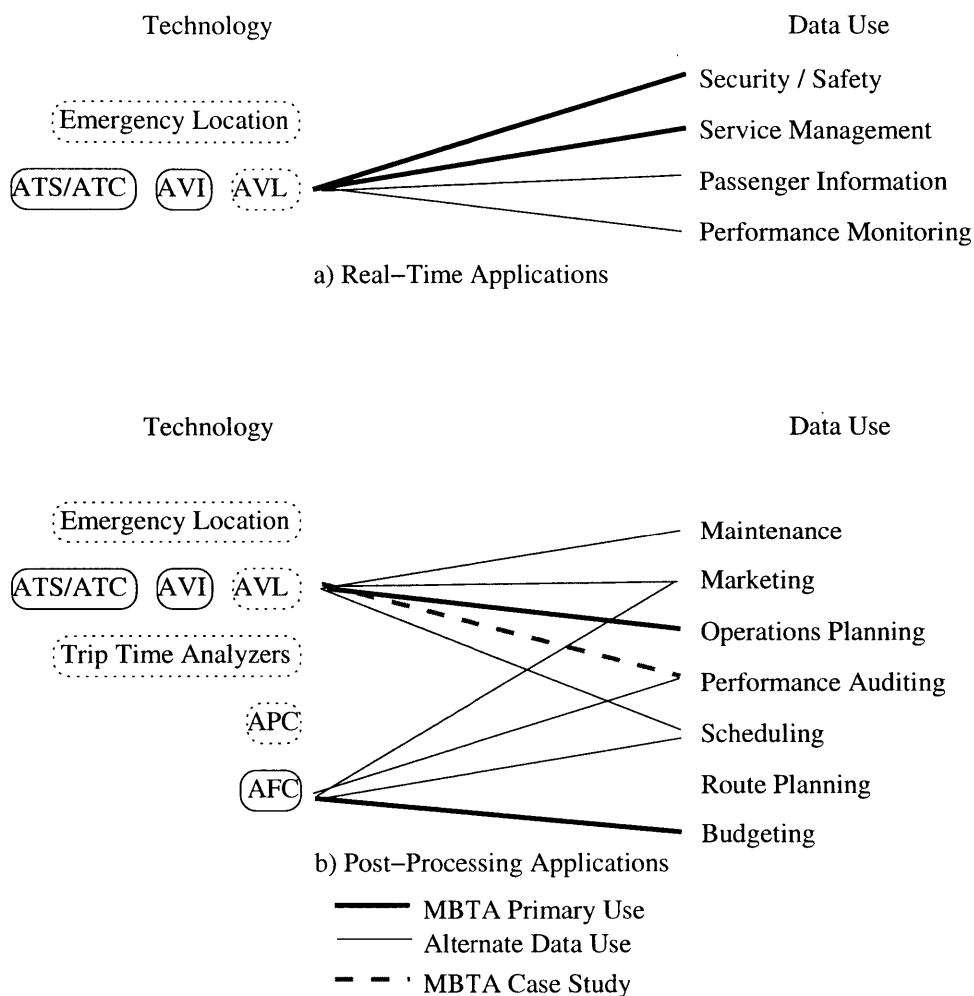


Figure 5-2: MBTA Use of Agency Technology

Figure 5-2 illustrates both current technology and planned future technology available at the MBTA, as well as the applications to which this technology is currently applied, or will be employed when installed. There are a number of options for providing extended capability for use of data to the Rail division of the agency using these sources of data.

5.1.3 Agency Needs and Opportunities

For each of the needs outlined in Section 2.1 this section identifies the degree to which they are currently satisfied and identifies potential improvements in performance pos-

sible given the particular sources of extensive data currently available and soon to be available to the agency.

Security / Safety

The MBTA rail division already fully utilizes its available train position data to calculate and display locations in the Operations Control Center for all trains on the Red, Orange and Blue Lines, as well as some indication of position on the Green Line. Any further advances in security would require additional infrastructure.

Service Management

The MBTA uses train position data for coordinated Service Management on the Red, Orange, Blue and Green Lines. Train position data is presented to individual controllers on desk-mounted CRTs as well as on large-screen data projectors for system-wide oversight and coordination. While controllers already have a wealth of data available to them, there are a number of opportunities to further improve service management. These opportunities include controller displays that summarize operations in novel ways, tools to aid with dispatch control, as well as the distribution of real-time train position displays to field supervisors.

Alternate line summary position displays, denoted application “sm1”, could include a real-time display of vehicle progress over time in dynamic time-space diagrams for controllers. Such a presentation of position of all vehicles along a route from ATC data would show developing gaps or train bunches at a glance. The primary benefit of this presentation over a simple snapshot of vehicle position in time is that the time-space diagram illustrates the typical segment running time. While experienced controllers are already aware of the time-varying typical run times, a more information-rich display of position would facilitate maintaining regular headways. The current display visually encourages controllers to evenly space vehicles in distance, however since run time varies between parts of the route this will not always lead to even headways. If the vehicle position display is adjusted to provide segment lengths proportional to the typical running time then this could provide the proper

information, but this would need re-adjusted throughout the day as typical segment level running times change.

Dispatch control is vital to provision of reliable service. An automated dispatch control system, application “sm2”, has the potential to improve service quality by reducing initial headway variance through an automated indication of the best time to leave the terminal to provide even headways. This can be especially important in adjusting departures to coordinate interleaving between branches of the Red Line. Such an application would use real-time ATC derived position of all vehicles in the terminal area. The Green Line stands to benefit greatly from automated dispatch control, however terminal operations are not adequately instrumented by the Green Line AVI detectors to take advantage of automated dispatch control. A pilot project for Red Line dispatch control has been tried on a portion of one branch of the line and found to provide dispatch time recommendations that agree with decisions of experienced controllers. An extension of this system to cover the full Red Line could be very beneficial for headway regularity.

As mentioned in Section 5.1.1, field supervisors are responsible for the provision of regular service on the Green line. The information to which these supervisors have access is presently very limited on their own line, and is even more limited with respect to the position of vehicles on other lines. Wong [42] determined that dramatic reductions in passenger waiting time are possible through coordination between the Green and Red line at Park Street, especially through holds of Green Line Westbound trains. Building on this prior research, application “sm3” concerns itself with the provision of a display of Red Line position information to Park Street Green Line supervisors. This application is primarily concerned with display of heavy rail ATC data, in this case the Red Line, to field supervisory staff. Since this station is such an important transfer point, it is also an ideal location to provide a simple automated indicator, such as a “Hold” light, to alert a supervisor to an especially beneficial hold for an arriving Red Line train.

Passenger Information

As described above, the MBTA Rail network has four primary light and heavy rail lines, two of which branch, providing service to many dozens of underground and at-grade stops. In Boston's non-grid network of urban centers provision of passenger information is especially challenging. The MBTA provides a great deal of passenger information to help customers use the system, but the use of automatically collected data has the potential to further improve the transit user experience. Passenger information needs are split up into pre-trip, at-stop and en-route needs. First is pre-trip planning; this is important to best utilize rail to rail transfers and the bus feeder network. These pre-trip needs are addressed using traditional paper maps, web-based maps and a trip planner, as well as a Palm handheld-based trip planner which includes detailed schedule information for Rail as well as Bus, Ferry and Commuter Rail lines.

The web and Palm-based service provide only static schedule information. They do not provide information based on the actual state of service. There is a web page where notices are posted of large-scale disruptions in service, however that is not integrated into either the web or Palm-based trip planner. The MBTA does not take historic measures of service variation into account when suggesting departure times and designing connections. This run-time variability and schedule adherence information may be derived from the current train position data that is already available from the MBTA ATC and AVI systems, application "pi1". An advanced trip planner which uses already available, real-time train-tracking information could improve user's ability to, with a given (say 95%) confidence, arrive at a destination by an appointed time.

Another application, "pi2", involves the use of typical segment-level run times in combination with real-time ATC or AVI train position data to provide a web-based automated reminder of when to leave. An initial implementation may be passive, providing information for a user to judge when to leave, however optimally the agency application could take an active role in warning when to depart. An active system also requires either an automated calculation of the station access time or simply

an entry for how many seconds warning a user wants, which could be adjusted to accommodate the individual user's walking speed and other factors influencing access time. Both active and passive systems could improve the attractiveness of transit by reducing time spent leaving early to assure an on-time arrival at a destination despite possible system delays. While last-minute system delays will of course sometimes happen, information about the current state of operations will reduce their impact and uncertainty.

The agency has 60 heavy rail and underground light rail stations. It has another two dozen at-grade light rail stations, as well as a large number of at-grade light rail stops. Addressing the current state of in-station information, all MBTA stations are clearly signed and have system maps, and a number have more detailed area maps including local bus connections as well as street names. The frequent at-grade light rail stops are often less well signed, with less information available to the user. Rail schedule information is only available at a small number of stations; most do not even provide expected headway for the time period. However the Palm-based tool for trip-planning can provide some users some in-station schedule and trip-planning information. In station next-train announcements are often given over the station PA system, particularly if there is a disruption.

The greatest opportunity to improve passenger information in the rail system is to provide estimates of time until next train arrival on variable message signs, as data of adequate quality is currently collected for the heavy rail routes as well as the Green Line. [Application "pi3".] Time until next train arrival at a particular point can be calculated by combining ATC or AVI derived position along with current or typical running time information. A few stations have variable message signs installed, however they display static announcements and are not updated with vehicle arrival information, and these signs are not yet connected to a central computer.

On-vehicle information is primarily limited to station announcements and a line map over each door. The newer Red Line vehicles have audible and visible automated stop and destination announcements, however there are no plans currently in place to update the still-in-operation older Red Line vehicles which require manual audible

stop announcements. An on-vehicle improvement currently possible works in concert with the MBTA's recent adaptation of its subway tunnels to allow cellular telephone use while underground. By adapting their website to work with the wireless-web browsers available on newer cellular telephones, the T could provide improved on-vehicle information. This application, "pi4", concerns itself solely with the adaptation of the current website to provide pages in the Wireless Application Protocol [WAP] standard format used by wireless web browsers. It will include whatever information is on the current website, specifically schedule and route information, and potentially real-time position information of connecting vehicles using rail position data to help the user dynamically choose a minimum time route.

Maintenance

There are opportunities to improve rail maintenance at the MBTA by using currently collected ATC and AVI data to audit indicators of vehicle and detector health. One such opportunity is to automatically identify Green line AVI detectors that regularly fail to identify passing vehicles. [Application "maint1".] This would help the Green Line controller, and other rail controllers who rely in Green Line position information for system coordination.

Another opportunity is to use ATC data to automatically identify individual vehicles with indications of malfunctions. [Application "maint2"] For example, by observing dwell times for all vehicles at all en-route stations and comparing those dwell times to average dwell times for that type of vehicle at that station at that time of day, sets of rail cars [aka consists] with consistently higher than average dwell times could be singled out for inspection of proper door function. In this case no information about passenger load or demand is needed since vehicles are compared to one another over different, comparable, days, assuming there is variation on which particular cars make a particular run on any given day.

Lastly, a small hardware investment in in-vehicle or near-alignment sensors which relay performance to a central computer could be very worthwhile if combined with vehicle position data. For example, information from near-alignment vibration and

sound pressure level [SPL] sensors could be combined with ATC information on the position of individual consists to provide SPL measurements on particular vehicles. [Application “maint3”.] In addition to immediately identifying unusually loud vehicles, these measurements can be tracked over time to allow identification of vehicle performance degradation.

Marketing

Marketing effectiveness can be improved by a stronger focus on specific target audiences. By extending the web-based trip-planner to provide an option for users to submit their contact information and be contacted for changes in service that impact them. While some information is collected in registration for the Palm-based trip planner, that information is collected by a separate company – Second Kiss Wireless. By considering this registration data as an agency source of automatically collected information, an example use of this, application “mark1”, would be for marketing extended Blue Line service hours or a new connecting bus route to Logan Airport if they’ve either explicitly requested updates on Airport issues or have previously requested trip information for connecting to the Airport. Another use would be to advertise a new trip-planner feature or interface and use these registered users for testing before release to the general public.

Operations Analysis

A wealth of information on operations is presently available from the high-quality heavy rail ATC data. Line capacity can be estimated, application “anal1” using the variability in the current operations to identify current peak throughput as well as identification of bottlenecks, such as a terminals, interlockings or stations with long dwell times. Green Line AVI data can be used [application “anal2”] to identify variations in performance between individual operators in a manner similar to the CTA Bus case study. Coordination between lines can be studied [application “anal3”] using an integrated approach to identify off-peak system performance improvements through line coordination. Also, one can study the possibility of peak-period improvements

through deliberately avoiding simultaneous station arrivals in cases where the transferring passengers often slow down the through traffic greatly. Circumstances can be identified where the added convenience to transferring passengers does not justify the dwell time penalty to through riders.

Using the forthcoming AFC data, analysis integrating ridership and train position is possible, application “anal4”. This data may be used to identify portions of routes with inadequate service as well as to observe changes in demand over time. A thorough analysis of the passenger impacts of terminal operations, application “anal5”, using detailed information about track blocks and minimum safe distances, as well as inbound and outbound demand can lead to suggestions for best practices in terminal operations. Such an analysis may recommend changes in scheduled layover time or minimizing passenger impacts by holding a train at the previous station which will be subject to terminal congestion.

Performance Monitoring and Auditing

Performance monitoring and performance auditing are closely related activities but on different time scales. Both performance monitoring and auditing are done informally at the MBTA by those responsible for service management – light and heavy rail controllers and light rail field supervisors. Automated tools to aid in performance monitoring might combine measures of service performance in real-time from train position data to provide summaries of the quality of service being provided. [Application “pmon1”.] Performance auditing provides these numbers after the fact for identification of highly effective and less effective or counterproductive service management strategies. [Application “paul1”.] While averages are useful, it’s also often the case that there is a threshold effect with passengers. Passengers may be permanently lost to transit by repeated instances of denied boardings or extreme waits. It is useful to include performance measures that capture these effects, such as percentages of passengers experiencing service worse than a set threshold.

As discussed in 2.1, the primary data product needed for performance monitoring and auditing of a high frequency heavy rail line is the position along the route for

all trains. This allows calculation of headway adherence and variation. This position information is available from the ATC system for the heavy rail line, and data from the Green Line AVI system is adequate for such calculations on portions of the line. Augmenting this position information with demand information can provide a more customer-oriented summary of daily operations. For example the passenger impacts of interventions, such as short-turns can only be estimated using passenger demand information to estimate expected vehicle load. A system which incorporates demand information, applications “pmon2” and “paul2”, will weigh stations with high demand more heavily than stations with low demand when wait times are averaged. This will tend to reward operations strategies which serve the largest number of passengers well.

As mentioned in Section 5.1.1 automatically collected demand information is presently unavailable. However, manually collected passenger demand data from 1995 are available, showing demand per station in 15 minute intervals. AFC data will provide more frequent, possibly even real-time, updates of this demand information, and will also be able to show details not manually collected such as variability between weekdays as well as demand information for weekends. But this AFC data will have ambiguity in destination for many stations as the inbound and outbound platforms share fare collection equipment. For threshold values, schedule information is also needed to evaluate actual versus expected performance, such as for the percentage of passengers having to wait more than two scheduled headways.

Scheduling

MBTA Rail scheduling uses point checks from CTPS, as mentioned in Section 5.1.1 to audit train running time and load to assure accuracy of schedules and adherence to load standards provided in the MBTA Service Standards [23]. However, such checks are infrequent. Since the MBTA uses Hastus [12] for scheduling, they have the option to purchase a companion product: Hastus-ATP that optimizes segment level running times based on input from sources of automatic vehicle location data. However, this sort of segment-level run-time data is useful in other parts of the agency, such as

Passenger Information, and it may be worthwhile to calculate such run time figures in-house from available ATC and AVI data. [Application “sched1”.]

An immediate use of such data is to improve converging branch coordination on the Green and Red Lines. [Application “sched2”.] Some work has already been done on this topic [33], but further work is needed to provide regular headways in the trunk portion, as well as to reduce needless, predictable in-vehicle passenger delay at points where the lines converge. Data from automatically collected sources which will provide measures of demand, such as the upcoming AFC system, can be used to audit ridership and conformance with MBTA Load standards. [Application “sched3”.] Automatically collected data is especially useful in the off-peak and on weekends since they often are checked the least frequently

Route Planning

Rail route planning is infrequently needed, however route planning for rail feeder buses is a frequent need. As with Scheduling, automatically collected sources of demand data, for off-peak and weekends, can help determining whether a new feeder route will be successful, since quality weekend service can be important for maintaining a transit user base. The AFC system will provide such data, when it is installed, however in the meantime some demand information can be inferred from requests to the trip-planning portion of the MBTA website. [Application “plan1”.] While not all requests are from actual or potential users, and some users make multiple requests for a given trip, there are well-established techniques for identifying and distinguishing individual web users. [Application “plan2”.] This can be a source of highly desirable Origin Destination demand to augment such information collected from users through more traditional survey methods.

Budgeting

There are often benefits to the budget from many of the previously listed uses of automatically collected data, for example a schedule set using automatically collected data may reduce wasted time and/or may reduce overtime pay incurred when an infeasible

ble schedule consistently requires operator overtime. However these only indirectly improve the MBTA's ability to budget.

5.1.4 Intermediate Data Products

This section first summarizes the information needs of each of the applications identified in Section 5.1.3. From this it identifies the commonly used intermediate data products. Next it reviews the difficulties in fulfilling each intermediate data requirement. This starts with information derived from extensive position data – ATC- and AVI- based data in this case. Next it addresses information derived from extensive data on demand – primarily from the forthcoming AFC system. Last is information only available through combinations of extensive data sources.

Listing of Application Information Needs

Using Figure 3-2 as a guide, this subsection summarizes the information needed for each application in order to identify the most commonly used intermediate data products. Once an intermediate data products is found to be used in multiple applications, it is assigned an identifying letter. This letter is used later in the list in the place of detailed information on how to generate that intermediate product.

- Alternate Line Progress Displays [sm1]
 - Real-time position of all vehicles on route [P]
 - * Real-time ATC/AVI Data
 - * Position of each ATC/AVI block/detector
- ATC/AVI Data for Dispatch Control [sm2]
 - Real-time position of all vehicles on route [P]
- ATC Data for Field Supervisors [sm3]
 - Real-time position of all vehicles on route [P]
- Trip-planner with reliability info [pi1]

- Segment-level run-time / reliability [R]
 - * History of position of all vehicles on route [H]
 - History of ATC/AVI Data
 - Position of each ATC/AVI block/detector
- Schedule adherence [S]
 - * Schedule information
 - * History of position of all vehicles on route [H]
- Web trip-planner with real-time info [pi2]
 - Latest possible time to begin journey
 - * Real-time position of all vehicles [P]
 - * Segment-level run-time / reliability [R]
- Real-time info on Variable Message Signs [pi3]
 - Time until next train arrival at a station
 - * Real-time position of all vehicles [P]
 - * Segment-level run-time / reliability [R]
- WAP Enabling Existing Trip Planner [pi4]
 - Generation of WAP pages
- AVI Detector Health [maint1]
 - Malfunctioning detector list
 - * History of AVI position data [H]
- ATC Data Mining for Maintenance [maint2]
 - History of ATC position data [H]
 - Individual station/train dwell times [D]
 - * History of ATC position data [H]
 - * Position of stations / base dwell time.
- Noise Measurement Using ATC Data [maint3]
 - Per-vehicle noise measurements

- * History of ATC position data [H]
 - * Noise level data
- Marketing to Registered Web Users [mark1]
 - Indexing of website preference logs
- Line Capacity Estimate [anal1]
 - History of ATC position data [H]
 - Individual station/train dwell times [D]
- Operator Variability Analysis [anal2]
 - History of AVI position data [H]
 - Operator information
- System Line Coordination [anal3]
 - History of AVI position data [H]
 - History of multiple lines' ATC position data [H]
 - Load profiles / Station-level demand data [L]
 - * Passenger Counts
 - * AFC Data for Updates
 - Individual station/train dwell times [D]
- Dynamic Load Estimates [anal4]
 - Predicted Vehicle Load
 - * Real-time position of all vehicles [P]
 - * Load profiles / Station-level demand data [L]
- ATC Terminal Operations Analysis [anal5]
 - History of ATC position data [H]
 - Load profiles / Station-level demand data [L]
- Performance Auditing Phase 1 [paud1]
 - Vehicle-based quality of service measures
 - * History of position of all vehicles [H]

- Performance Monitoring Phase 1 [pmon1]
 - Vehicle-based quality of service measures
 - * Real-time position of all vehicles [P]
- Performance Auditing Phase 2 [paud2]
 - Passenger-based quality of service measures
 - * History of position of all vehicles [H]
 - * Load Profiles / Station-level demand data [L]
- Performance Monitoring Phase 2 [pmon2]
 - Passenger-based quality of service measures
 - * Real-time position of all vehicles [P]
 - * Load Profiles / Station-level demand data [L]
- Refine Scheduled Run Times [sched1]
 - Segment-level run-time / reliability [R]
- Improved Branch Schedule Coordination [sched2]
 - Segment-level run-time / reliability [R]
 - Schedule Adherence [S]
 - Schedule
- Load Profile Estimation From AFC [sched3]
 - Load profiles / Station-level demand data [L]
- Measures of Zonal Demand [plan1]
 - Zonal demand
 - * AFC linked boardings by location
- O-D Figures from Trip-Planer [plan2]
 - O-D demand
 - * Trip-planning website requests

Application Description [ID]	P	H	R	D	S	L
Alternate Line Progress Displays [sm1]	P					
ATC/AVI Data for Dispatch Control [sm2]	P					
ATC Data for Field Supervisors [sm3]	P					
Trip-planner with reliability info [pi1]	P	h	R		S	
Web trip-planner with real-time info [pi2]	P	h	R			
Real-time info on Variable Message Signs [pi3]	P	h	R			
WAP Enabling Existing Trip Planner [pi4]						
AVI Detector Health [maint1]		H				
ATC Data Mining for Maintenance [maint2]		H		D		
Noise Measurement Using ATC Data [maint3]		H				
Marketing to Registered Web Users [mark1]						
Line Capacity Estimate [anal1]		H		D		
Operator Variability Analysis [anal2]		H				
System Line Coordination [anal3]		H		D		L
Dynamic Load Estimates [anal4]	P					L
ATC Terminal Operations Analysis [anal5]		H				L
Performance Auditing Phase 1 [paud1]		H				
Performance Monitoring Phase 1 [pmon1]	P					
Performance Auditing Phase 2 [paud2]		H				L
Performance Monitoring Phase 2 [pmon2]	P					L
Refine Scheduled Run Times [sched1]		h	R			
Improved Branch Schedule Coordination [sched2]		h	R		S	
Load Profile Estimation From AFC [sched3]						L
Measures of Zonal Demand [plan1]						
OD Figures from Trip-Planer [plan2]						

- P Real-time position of all vehicles on route
 H History of position of all vehicles on route
 h History of position of all vehicles on route needed indirectly
 R Segment-level run-time / reliability
 D Individual station/train dwell times
 S Schedule adherence
 L Load profiles / Station-level demand data

Table 5.1: MBTA Application Shared Information Needs Summary

Application Information Needs Summary

These letters are also used in Table 5.1 which summarizes the dependence of the listed applications on the six intermediate data products identified with a letter in braces in the preceding listing of application information needs. Applications which use [R] or [S] also use historical data indirectly, and this is made explicit in the table with a lowercase [h] to emphasize the importance of the underlying data source.

Intermediate Data Product Details

This section addresses the challenges involved in producing each of the six shared intermediate data products identified above, as listed in the key to Table 5.1.

History of position of all vehicles on route. [H] The position of vehicles on the heavy rail ATC system are logged in terms of track occupancy and track un-occupancy events. The system provides a track identifier [eg “22551T DAVIS SQUARE”] as well as the car number of the lead car in the consist. While the track un-occupancy records can give higher spatial resolution, this introduces a dependency on train length; position is best determined by track segment-on records alone, which is used to determine the time of arrival at a position along a route via a fixed mapping from each track segment to the distance from a designated terminal. Mapping from this track identifier to position along route just requires a table look-up of the position along route and possible a branch identifier which maps to each track identifier that may be returned. [eg a table showing “22551T DAVIS SQUARE” is 8635 units of distance away from the Alewife terminal on track 2 of the main line.] The track-on information is integrated with train-tracking information in an agency train-tracking database to provide the scheduled trip identifier for this route. Queries can be constructed to provide an entire days’ worth of tracking data, as well as to provide consist information from “train_sheets”. The main difficulties in this step are constructing the database query and collecting enough historical data to provide an adequate number of runs over a suitable period of time.

The database only provides the identification of the lead car in a consist; to be

able to link inbound and outbound trips, train consist information is needed. This can be inferred from the occupancy history – by keeping track of distinct trains – however consist information is also readily available from the database.

Green Line AVI data has traditionally been more difficult to access than heavy rail ATC data, although efforts have been made to bring this data into the centralized database. At the time of this agency application, Green Line data was only available from a mainframe data collection system as a text file listing of detections. Any use of this data in this format requires information on detector locations in order to map to distance along a route, and identify the branch involved.

Summary of data sources for position along route for all vehicles:

- Track segment-on data on entire rail line.
- Train consist information for logically connecting inbound and outbound trips.
- Position along route of each track segment.

Real-time position of all vehicles on route. [P] The AVI and ATC based position data above can be accessed with very little delay – in near real-time – by an application running on the MBTA network. Such clients of the data must be trusted to not adversely effect system operations; the database joins which are used to provide train tracking data for historical data access use too much resources to be done frequently or during the peak. Programs to access this real-time data may only be written by individuals with extensive knowledge of the real-time system constraints. Train-tracking data is a critical part of the service management and security functions at the MBTA. This lack of distribution is an obstacle to any system requiring this data; in order to facilitate a range of future uses an intermediate data distribution machine may be required.

Segment-level run-time / reliability. [R] The calculation of segment-level run times is trivial, simply taking the difference in time between any two points along a route. Taking the average of multiple of these run-times for a given segment provides

the required segment-level run-time. The variance of the numbers gives an indicator of the confidence that a train will run that segment within a certain window of the average run time provided. Unfortunately these variances are not independent from one track segment to another, so reliability information must be calculated separately for each pair of points of interest, such as between all stations within a certain number of stops of one another.

Another difficulty in this step is to split this information up by time period and by day to reduce the variability and be able to provide better estimates. Often with automatically collected data sources the quantity of data used is beyond that required for the desired level of confidence in the figures provided. However, as more segregation of days and time periods is done, it is important to use a statistics reference to assure that run-time and variance figures calculated are representative and precise.

Individual station/train dwell times. [D] By identifying the ATC segment containing each station, either dynamically or from static sources of information, it is possible to create a proxy for vehicle dwell time from a position history. [H] This can be investigated manually to identify typical acceleration and deceleration times for each station. Alternately the minimum dwell time, as measured in late evening service could be subtracted from each dwell time, to give the variable component based on passenger demand. Another difficulty is that this will be distorted by control interventions. This can be mitigated by primarily observing vehicles at times when interventions are not common.

Schedule adherence. [S] Schedule adherence calculations require vehicle position over time as well as the corresponding schedule information. As mentioned above, the database access providing historical train tracking data also has a scheduled trip identifier. This provides an indication in the “train_sheet” file of the scheduled versus actual departure and arrival time of the train. The MBTA does not use scheduled times at intermediate points along routes, however they are likely used

internally for schedule creation and available this way, or they may be estimated using the available scheduled departure time and segment-level run times derived above.

Load Profiles / Station-Level Demand Data [L] As addressed in Section 5.1.1, presently the most detailed demand information is available from passenger count surveys although by June 2005 an AFC system capable of providing centralized logging of actual system usage will be installed. The survey, last conducted in 1995, observes platform entry and exit for each station at a high temporal resolution, with information on the demand in fifteen minute time windows to show demand variation over time. However the data is showing its age and agency users are skeptical of its validity. It is possible to recalibrate this data with more recent information, however the only more recent system wide stop-level data available is from 1997, at lower temporal resolution, in a collection primarily focused on transfers.

The AFC system will be an entry-only system, so it alone will not measure passenger flow comprehensively enough to provide load profiles. There is also ambiguity in destination for counts in many stations, given that a turnstile is shared between inbound and outbound platforms. However a number of techniques exist to estimate this information using AFC boardings. One such technique is to infer alighting demand by linking passenger trips and assuming that the location of the next use of their fare card is the same location as where their previous journey ended. [9] However this relies on people using fare-cards and not traveling by another mode in between card use. Given an updated survey during AFC collection, the fare-card usage rate and any differences between fare-card user and non-fare card user behavior can be accommodated. If passenger exits through turnstiles are timestamped and counted, the timestamps can be compared with train arrival timestamps to estimate the direction of travel of a particular group of users leaving a station.

5.1.5 Application Challenges

This section reviews the challenges to implementation of each application identified in Section 5.1.3. This includes identification of difficult algorithms, challenging information distribution, unusually large quantities of data required, or special-purpose combinations of data sources. This section also references the difficulty in intermediate data product creation addressed in the preceding section. All of these factors give a sense for the application complexity, cost and time required to completion.

Alternate Line Progress Displays [sm1]

Providing alternate real-time displays of position data for operators, such as a dynamic time-space diagrams for controllers would require the real-time position of all vehicles on route [P] intermediate data product, along with other static information such as the distance-along line of each station and crossover to provide a context for the visualization of vehicle movement. The real-time vehicle position data is available, but as discussed in Section 5.1.4, it is somewhat difficult to access due to the risk of impacting operations. The static information on stations and crossovers is also available, but some manipulation would be required to get it into a format usable for display.

The actual presentation of the data can be done using freely available graphics plotting; several exist which are fast enough for once per second display updates without adding noticeably to machine load. Once the decision on which new display options is made – such as to provide rolling time-space diagrams or a progress snapshot with inter-segment distances proportional to typical running times for that time of day – the implementation of the display is straightforward.

ATC/AVI Data for Dispatch Control [sm2]

Extending the pilot Red Line Braintree branch Dispatch Control System [DCS] to provide dispatch recommendations for the remaining two Red Line terminals would require more vehicle position data [P] as well as a summary of current practices in

a way that can be expressed in logic to provide dispatch time recommendations that agree with decisions of experienced controllers, and aid in generation of even dispatch headways. The understanding of current practice could be derived from interviews with controllers and optionally analysis of historical data. As with other tasks requiring real-time access to vehicle position data, this program must be trusted to interact benignly with the live system, which will make it more difficult to implement.

ATC Data for Field Supervisors [sm3]

The provision of Red Line position information to Green line field supervisors presents novel challenges. One is the deployment of computer hardware in an underground rail tunnel, the steel dust and moisture providing an unusually harsh environment. Harsh-environment displays are readily available, but they are expensive. The distribution of the information will involve hardwiring data, and possibly a hold indicator sign. As discussed in Section 5.1.4, the system access to extraction of real-time vehicle position data [P] presents some difficulties, but extraction of position itself is straightforward, as is the algorithm for Green Line holding.

Trip-planner with reliability info [pi1]

To extend the web-based trip planner to include realistic line reliability information in its routing will likely be difficult. The individual data products that drive it are run-times and reliability [R] and schedule adherence [S]. The reliability numbers are straightforward to calculate, but require a large database of previous trips. The schedule adherence figures may or may not be needed based on the line frequency and time of day, but in cases where they are needed the main difficulty is obtaining schedule information; the combination of archived position information and schedule information should not be difficult since the scheduled trip id is available from the database. As the planner was not written in-house, any attempt to add this information in the planner will be through the vendor of the route-planning software. Even with schedule and reliability information readily available, the vendor's price to add this non-standard feature may be prohibitive.

Web trip-planner with real-time info [pi2]

In addition to the difficulties experienced in the previous application [pi1], of expected run times [R] and familiarity with extending a trip planning tool, this application adds the complexity of real-time position information [P] extraction and integration with the trip-planning system. The actual calculation of the latest possible time to begin a journey is also not trivial, especially in the case where the expected vehicle to be used will need to turn around at a terminal before returning to serve the user.

Real-time info on Variable Message Signs [pi3]

Providing train arrival message signs in stations faces a number of challenges. The first is the actual collection of information needed for the product, both the available but difficult-to-access real-time position [P], and the straightforward but data intensive segment level run time information. [R] Once this data is available it is relatively easy to combine this information in a prediction of vehicle arrival. The communications with the Variable Message Signs currently installed in a couple stations is the greatest difficulty as communications hardware and wiring is required.

WAP Enabling Existing Trip Planner [pi4]

The adaptation of the MBTA website to allow easy access from web-enabled phones does not require any new hardware or data collection. It is a simple extension of the current web-page design to generate Wireless Application Protocol [WAP] formatted pages, suitable for small text-only handheld devices. Second Kiss Wireless, the providers of Palm-based schedule information has a ready-made option to provide information in this format.

AVI Detector Health [maint1]

In order to produce a malfunctioning detector list, only very simple information is required. A short – one day will do – history of AVI position data [H] is all that is required. The challenge is in train-tracking of Green Line trains, identifying the

normal sequence of sensor detections, and then identifying situations where a detector is passed but no record is present for a train detection. This can be done visually with human interaction, by presenting a plot showing detectors as nodes with vehicle progress from one detector directly to another showing up as a directed arc between nodes, however an automated health monitoring system has the potential to catch errors more quickly.

ATC Data Mining for Maintenance [maint2]

Using extensive ATC position history data [H] to automatically identify individual vehicles with indications of malfunctions may require processing a great deal of data in order to identify out-of-the-ordinary behavior. It also requires some inspiration as to what sort of behavior is normal variation, and what behavior – such as long dwell time or run time – indicates an impending problem. One concrete example is the examination of dwell times, [D] as discussed in Section 5.1.4, where given a suitably long period of observations interventions should be distributed over different vehicles and inter-vehicle variation may become apparent. Observing changes in dwell [D] or run-time [R] over time can also help to identify problem vehicles. Ideally this could be trained by information on actual vehicle failures and preventative maintenance, to help identify precursors to failure, and figures which are improved by maintenance.

Noise Measurement Using ATC Data [maint3]

A specific indicator of vehicle health is vehicle noise and vibration; in order to instrument this with per-vehicle noise measurements, added hardware is required to log vehicle noise performance. Vehicle proximity information from ATC position history data [H], can be compared to identify individual vehicles as sources of noise and vibration. Background noise sources will alter the readings, but median values should be consistent over time. Similar to the previous application, any change in these figures over time can be informative, as well as any change observed just before a failure or just after preventative maintenance will show the relation of noise and vibration to vehicle health.

Marketing to Registered Web Users [mark1]

A simple extension of the web-based trip-planner to provide an option for users to submit their contact information can provide a useful set of marketing information. The changes to the web site are minor, requiring a new straightforward form to allow users to express their preferences. The collection of form-submitted data and processing of this data to drive new marketing strategies is also simple. The greatest challenge is to provide the right options on the transit user preference form to get sufficient information from users to direct new marketing campaigns. Using website cookies to track individual user search requests, and the use of these search requests to target marketing is also possible, but requires more effort to design as well as to use the travel demand data well for marketing appropriate services to the user.

Line Capacity Estimate [anal1]

Estimation of line capacity can be done using an extensive history of ATC derived position data [H]. The primary challenge in this analysis is the determination of both current system bottlenecks, and which future bottlenecks will emerge when the worst are removed. This can be identified through observations on minimum headway approaching a station and minimum headway departing, as well as by observing dwell [D] times at different times of day over different parts of the line.

Operator Variability Analysis [anal2]

Combination of historical position data [H] and operator information can be straightforward or difficult depending on the manner in which operator information is stored. Obtaining the operator information in an electronic format with scheduled trip numbers, for example would make this combination easy; only obtaining a printout of operator pick sheets with run times and no identifying trip numbers would be quite difficult to use with the available position data. Once the data have been combined then analysis to identify operator variation is a straightforward comparison of, say, run times for particular operators compared to the average run-times [R].

System Line Coordination [anal3]

Coordination between multiple lines historic position data [H] and optionally load profiles [L] and typical dwell times [D] can be very difficult. The position data is easy to create, and reasonable approximations of the dwell times and load are created and used as described previously. The challenge is integrating the information to provide useful figures for passenger impacts of lack of coordination, and even ideal coordination schedules.

Dynamic Load Estimates [anal4]

Using the forthcoming AFC data with turnstile data transmitted to a central computer, analysis integrating ridership and train position is possible, even in real time. The steps taken to provide load profiles in Section 5.1.4 can be applied in real-time, however this is subject to the difficulties in obtaining real-time position [P] information described previously, as well as uncertainty about the exact format in which AFC data will be shared.

ATC Terminal Operations Analysis [anal5]

In combining detailed information about track geometries aimed at avoiding terminal congestion, the actual history of ATC position data [H] and optionally inbound and outbound demand from load profiles [L], a more rigorous treatment of terminal operations can provide figures for the passenger impacts of various train operations procedures. This is more difficult than most as it involves integration of a number of sources of data, but it does benefit from the similarity of passenger assignment required in this analysis and in the generation of load profiles.

Performance Auditing Phase 1 [paud1]

Providing vehicle-based quality of service measures is a fairly straightforward use of position history [H] data. Since the data is very reliable – all the trips are present – using this position data along with information on the position of all stations along

the route can identify the headways experienced at each stop by walking forward through time for trains at a particular station. The variance of headways arriving at all stations will be correlated with the passenger waiting experienced. However it will overestimate waiting for situations where headways are irregular for reverse-direction service and regular in the peak-direction. Since variance is not a very accessible figure, a threshold figure, such as the number of headways at stations which are greater than twice the scheduled headway can provide a simple, easy to understand, summary of the quality of service. This scheduled headway information is easy to find in public sources.

Performance Monitoring Phase 1 [pmon1]

Performance monitoring is very similar to performance auditing, but using real-time position data [P], and providing service quality estimates over a shorter period of time – for example the last half hour – than a service auditing tool. The algorithms are very similar, and display of a handful of numbers is not difficult; the primary challenge in this application beyond “paud1” is the added difficulty of obtaining real-time position [P] information.

Performance Auditing Phase 2 [paud2]

Passenger-based quality of service measures build off the headway information available from application “paud1”, combined with the passenger demand during each headway. This is possible using station-level demand data over time [L], identifying the different demand patterns in the different time periods. The assignment of passengers to headways is straightforward, with the one difficulty being passenger assignment on a branching route, given that survey data does not differentiate arrivals at trunk stations based on eventual destination. The main difficulty is obtaining data in a suitable format for automatic application. This provides not only the total number of passengers served, but also how long they had to wait for a train. AFC-derived data can be used, but it has even more ambiguity in terms of passenger destination than the manual counts as described in the generation of the load profile

/ demand data product [L]. This information can be summarized over time manually in a spreadsheet or automatically as part of auditing to provide a comparison from one day to the next.

Performance Monitoring Phase 2 [pmon2]

Similar to the relationship between “paud1” and “pmon1”, this application is very similar to “paud2”, except real-time position data [P], and providing service quality estimates over a shorter period of time than a service auditing tool. The primary challenge in this application is the added difficulty of obtaining real-time position [P] information. Also, when the AFC system is online, static load information [L] can be supplemented or replaced with dynamic information about actual demand patterns. The difficulty of this task is also in the collection of the demand data from the AFC system without impacting normal system performance.

Refine Scheduled Run Times [sched1]

The creation of segment-level run times [R] is described in Section 5.1.4 as they are used in a number of applications. The actual creation of the times and reliability figures from position history data [H] is straightforward, but the application of these figures in refining run times provides some challenge. The main difficulties are formatting the times to interface with the Hastus scheduling tool, and also the proper use of the full route run-time figures to provide an optimal recovery time that assures a certain percentage of vehicles will be able to start their next trips on time.

Improved Branch Schedule Coordination [sched2]

Creating a schedule which coordinates branch operations is a special case of the first Scheduling related application [sched1]. The difference in this case is that in addition to the segment-level run-time and reliability [R] figures needed, the schedule adherence [S] figure can be used to assure that the distribution of arriving vehicles will likely provide a vehicle interleaving which creates a minimal variation in headway in the trunked portion of the line. This is mostly a matter of changing schedule time

on one branch by a few minutes, however situations when incompatible headways on the joining branches provide more of a challenge because vehicles do not naturally interleave evenly.

Load Profile Estimation From AFC [sched3]

This application has a variety of other users. In creating station level demand data [L], this improves both the ability to schedule as well as simplifies a number of other projects. For details on this project see the Load profile [L] entry in Section 5.1.4.

Measures of Zonal Demand [plan1]

Mapping AFC boardings to their geographic location of origin is straightforward for rail stations since the boardings occur at one of a short list of known static locations. The difficulty of summarizing geographic regions is reduced if the agency already has a GIS system for evaluating expected demand for new services.

O-D Figures from Trip-Planer [plan2]

Using the O-D request from the current trip planning website requires no new data collection, just web log analysis. Implementing the techniques for identifying unique web users presents some challenge, but e-commerce and internet advertising have driven a great deal of progress in this area. Given records of individual route requests desired origin and destination information can be integrated into an agency GIS database of desired demand. This can be straightforward if agency information is already geocoded and available in electronic form. A random survey to assess how representative the web requests are, and of what segment of the population, is an important addition, but adds a great deal of cost, delay, and complexity to the utilization of this data source. Once it has been confirmed as useful it does continue to collect data without further costs, though.

5.1.6 Application Selection

This step of the MBTA Rail agency appraisal combines the information on application benefits in Section 5.1.3 and the information on costs in Section 5.1.5 and Section 5.1.4.

Application Description [ID]	Expected benefit	Expected complexity	Time frame
Alternate Line Progress Displays [sm1]	1	2	soon
ATC/AVI Data for Dispatch Control [sm2]	3	4	soon
ATC Data for Field Supervisors [sm3]	4	3	soon
Trip-planner with reliability info [pi1]	1	4	now
Web trip-planner with real-time info [pi2]	4	4	now
Real-time info on Variable Message Signs [pi3]	5	3	soon
WAP Enabling Existing Trip Planner [pi4]	3	3	now
AVI Detector Health [maint1]	2	2	now
ATC Data Mining for Maintenance [maint2]	2	3	now
Noise Measurement Using ATC Data [maint3]	3	4	soon
Marketing to Registered Web Users [mark1]	2	2	now
Line Capacity Estimate [anal1]	4	5	now
Operator Variability Analysis [anal2]	2	3	now
System Line Coordination [anal3]	2	3	now
Dynamic Load Estimates [anal4]	4	5	future
ATC Terminal Operations Analysis [anal5]	4	4	now
Performance Auditing Phase 1 [paud1]	3	1	now
Performance Monitoring Phase 1 [pmon1]	2	4	soon
Performance Auditing Phase 2 [paud2]	4	3	now
Performance Monitoring Phase 2 [pmon2]	4	5	soon
Refine Scheduled Run Times [sched1]	1	1	now
Improved Branch Schedule Coordination [sched2]	2	2	now
Load Profile Estimation From AFC [sched3]	3	4	future
Measures of Zonal Demand [plan1]	1	3	future
OD Figures from Trip-Planer [plan2]	2	3	now

Table 5.2: MBTA Application Choice Matrix

Table 5.2 was generated in the same way as Table 4.2 in the CTA case study. (See Section 4.1.6) In this case study the most used intermediate data product is position history information [H], used in 15 of the 24 projects. It is very simple to generate based on ATC and AVI records, so the evaluation is not redone for the case where this product is already available. Having easy access to the second most used data product: real-time position information [P] would reduce the complexity of 8 of the 24 of projects. However, all projects which simplify real-time data access require new hardware in the form of a position redistribution machine to isolate new applications from the live systems. Since this requires extra hardware it is put in the

'soon' category. In this case study only applications that may be finished in a short time are considered, so the difficulty evaluation is not repeated under the assumption of a real-time position redistribution machine. Eliminating all projects where the time frame is 'soon' and 'future' as well as projects with a benefit that is less than or equal to the estimated project difficulty narrows the field to only two. The two applications with the most benefits compared to cost are both performance auditing applications, "paud1" and "paud2". Since the second application subsumes the first, and it was possible to accomplish within the allotted time-frame, this passenger-centric performance auditing application is selected for trial implementation.

5.2 Agency Benefits

To complete this case study I implemented a working prototype of the passenger-centric service quality auditing application. The following discussion begins with information on the data products made available by this application which can be used by other applications. Next it describes the benefits of the service quality / performance auditing application.

As described in Section 5.1.4, the position of vehicles on the heavy rail lines can be determined from ATC system track occupancy records which are available in an agency database. In order to facilitate further use of this data a fellow student, Adam Rahbee, set up a mechanism which at 5:30am automatically generates a database dump of the previous day's data, and sends that raw file as an attachment in an email for future access. The daily dump contains data from 2:00am the previous day to 2:00am of the current day. The bulk of the email consists of the position dumps for each of the three heavy rail lines, of which there are over 100,000 records total for a given weekday. Each record gives the segment identifier, time of segment occupancy along with identity of the the lead car as well as the scheduled trip identifier. Auxiliary data files available are the consist information "train sheets" from each yard, as well as the status of the terminal start light which signals vehicle departure. These emails were redirected to a machine set up to automatically process the position data into

a format readily accessible to data processing and graphics plotting tools, for further use. Since the data is derived from hard-wired frequent signpost-type detection, it is trivial to map from block location to position on route. Since only the lead car of each consist is identified in the raw data dumps, a consist identifier is added to allow tracking of vehicles through a lay-over.

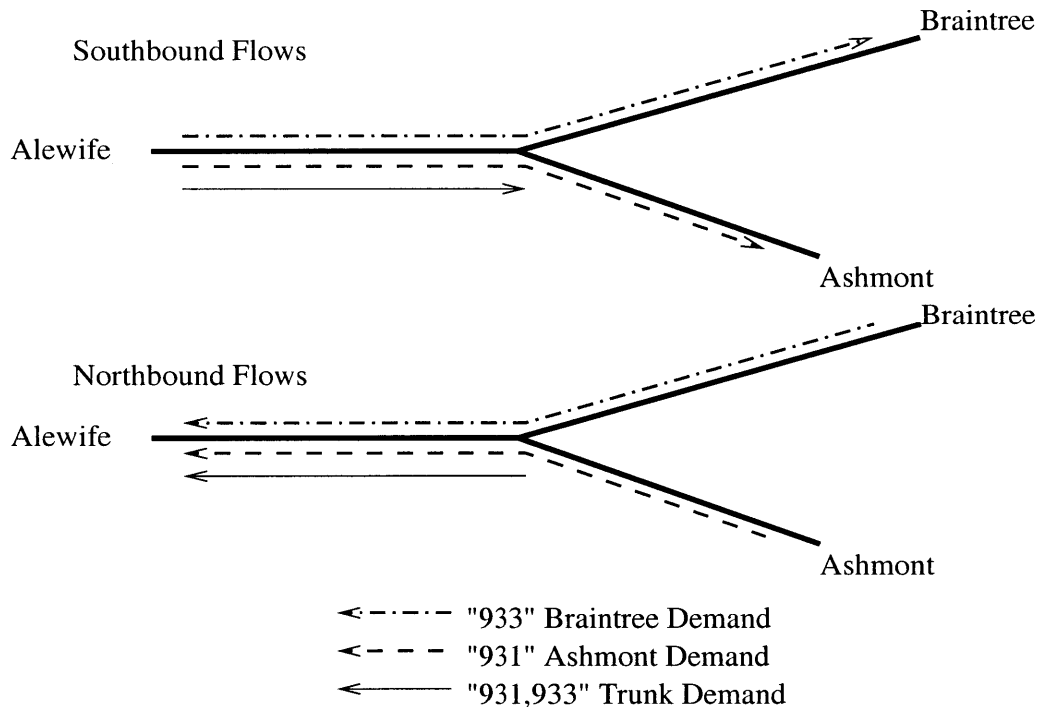


Figure 5-3: Schematic of MBTA Red Line.

Another data product of possible agency-wide use is an easily accessible form of train demand data. Since AFC-derived data will not be available in the short term, historical demand data was reformatted and processed to provide numbers usable for predicting vehicle loads. CTPS has passenger demand data for the MBTA heavy rail lines at a 15 minute resolution collected in 1995. While two heavy rail lines have no branches, the Red Line has a single branch, as illustrated in Figure 5-3. Passengers waiting for a Northbound train on the Braintree branch or for southbound train heading towards Braintree are in the 933 category, similarly for the Ashmont branch denoted 931, and the indifferent category, including all Northbound passengers when the lines are combined, as well as a fraction of Southbound traffic when the lines are split. In order to estimate passenger waiting times on this branching network,

the destination preference of arriving passenger is needed. Since this is not part of the available passenger demand data it was estimated based on the data collection's observed vehicle loads at the line split. Details of this calculation are in Section 5.3.1. This data is now available for use in other MBTA-related projects.

5.2.1 Service Quality Auditing Application

This section focuses on the use of ATC derived train position data, combined with historical demand data to provide service quality summaries from a user's perspective. Performance measures such as average passenger wait time, percentage of passengers waiting more than two scheduled headways, and other measures indicative of performance are calculated automatically nightly. These figures are sent using electronic mail to managers in the early AM to provide a summary of the previous day's operation, and to flag incidents that may be worth further investigation. A website is provided for further visual investigation of the previous days line performance, including informative views of terminal operations. Summaries for multiple days of operations are also provided to put the current day's performance in context, as well as to facilitate identification of poor and good performance. Another benefit of this project is the detailed information on the processing requirements of this prototype system, aiding in the specification of resources in the agency for a permanent installation in the agency.

The primary users of data from this system are the line managers, charged with maintaining the high performance of the heavy rail lines, though other operations analysis may also take advantage of the data. The system calculates a number of level of service figures from this combination of actual headway profiles and estimated demand values. One is average passenger waiting time, the percentage of passengers predicted to have waited less than one scheduled headway and the percentage of passengers waiting more than two scheduled headways. Another are the percentages of vehicles arriving at less than 1.5 scheduled headways or more than 2 scheduled headways. For more detailed analysis, graphical time-space diagrams of the line at various resolutions are needed. To aid in indexing, multi-day plots are also provided,

summarizing the above passenger waiting times and summaries of vehicle headway deviations.

An integral part of these estimates for branching routes is the split of vehicles and passengers based on destination. The vehicle destinations are provided in the database data dump, allowing simple calculation of headways on the trunk of the line for vehicles destined for individual branches as well as for the straightforward combined headway. Splitting demand based on destination is also done using historical data as described in Section 5.3.1. Having both vehicle position and demand split based on destination allows figures combining vehicle and passenger information, such as expected average passenger waiting, to take into consideration passenger destination.

Daily Service Quality Email

One component of the application is the automatic distribution of a performance email sent each night to myself and to the manager of rail operations at the MBTA. This email contains all the passenger and vehicle performance figures listed above, for each branch of each line, as well as a summary of unusual / exceptional events on each line. An excerpt of this is shown in Figure 5-4, identifying stations where trains dwelled excessively, or where trains took longer in transit than normal, or where a train was turned mid route. The lines indicating “delay” could be either an actual incident of a stuck door, or a non-cooperative passenger, etc., or they could be the result of a controller intervention to re-space the route. This system does not have any information on interventions to use to identify true disruptions. In general “slow” indicators are likely to be signs of incidents or congestion; in this case terminal congestion has delayed a train and its passengers just before the platform for over four minutes.

Another category of information in the daily email is a listing of a number of numerical performance measures. The first is the average waiting time based on the 1995 demand data. (See Figure 5-5.) For the branching route this arrival demand is split based on a per-station estimate of the the demand detailed in Section 5.3.1. For each train arrival at each station the number of passengers arriving as well as the

4/1/2002 16:04:55, 1860 delay 5:01 at ‘‘Park NB’’
 4/1/2002 19:16:43, 1600 slow (4:22) btwn TERMINAL and ‘‘Quincy Adams SB’’
 4/1/2002 19:57:49, 1822 delay 7:05 at ‘‘Shawmut SB’’
 4/1/2002 20:04:06, 1802 delay 9:22 at ‘‘Park NB’’
 4/1/2002 21:25:15, 1823 Short Turn btwn ‘‘Park SB’’ and ‘‘Charles/MGH SB’’

Figure 5-4: Red Line Exceptions excerpt for 4/1/02

Average waiting in seconds (number of passengers)				
RED Line	6:30/8:30	8:30/16:30	16:30/18:30	18:30/23:30
RED Line	181 (37688)	250 (98005)	189 (48788)	258 (30889)
933	215 (11155)	346 (17361)	249 (11330)	365 (4879)
931,933	144 (20767)	196 (65519)	150 (31403)	212 (22173)
931	252 (5766)	372 (15125)	275 (6056)	388 (3837)

Figure 5-5: Average Wait excerpt for 4/1/02

Branch Name	ID	Headway (minutes)		Ideal Avg Waiting (seconds)	
		Peak	Base/ Evening	Peak	Base/ Evening
Braintree	933	7	12	210	360
Trunk	931,933	4	6	120	180
Ashmont	931	8	12	240	360

Table 5.3: Red line scheduled headway and ideal average wait times

length of time each has waited is estimated assuming a uniform arrival rate within each 15 minute time period in the demand data. It also assumes that passengers board the first train that arrives which is heading to their destination – that vehicles are not over capacity.

Passenger waiting for each station and train arrival is combined to provide a figure which is presented split by route, by time period, and by passenger type. Ideally the average waiting time would be half the headway, which is summarized in Table 5.3. In the midday on the Braintree branch Figure 5-5 shows the service provided is better than the ideal for the published scheduled headway, indicating that trains are actually running more frequently than the published headway. Comparisons of these numbers from one day to the next are provided in the section on multi-day plots.

Percent of pax waiting less than 1 scheduled headway / more than 2				
RED Line	6:30/8:30	8:30/16:30	16:30/18:30	18:30/23:30
RED Line	86.9%/0.6%	86.5%/0.4%	82.7%/2.8%	84.7%/1.2%
933	90.5%/0.1%	91.0%/0.0%	83.9%/0.4%	90.5%/0.0%
931,933	83.3%/1.1%	84.7%/0.6%	81.8%/4.1%	82.8%/1.7%
931	93.2%/0.0%	89.1%/0.0%	84.8%/0.0%	88.2%/0.0%

Figure 5-6: Wait-Time Percentages excerpt for 4/1/02

Also provided are the expected percentages of passengers waiting less than one scheduled headway and the percentage of passengers waiting more than two scheduled headways. In a perfectly functioning system 100% would be in the first category and 0% in the latter. Sample output for the same period of time as above is shown in Figure 5-6. In this case you can see that a fairly large percentage of PM peak passengers using the trunk portion of the route had to wait for more than two scheduled headways. The scheduled headway used is the advertised headway for the time period, not the actual scheduled headway for vehicles around the headway in question.

There was some skepticism as to the validity of 1995 demand data for present day demand, so a vehicle headway based figure is included as well. In this figure, presented in Figure 5-7, is analogous but not directly comparable to the passenger wait figure in Figure 5-6. This figure shows the percentage of vehicles arriving at less than one and a half scheduled headways and the percentage arriving at greater than twice

Percent of vehicles at less than 1.5 scheduled headway / more than 2				
RED Line	6:30/8:30	8:30/16:30	16:30/18:30	18:30/23:30
RED Line	82.3%/4.9%	91.0%/1.6%	87.7%/4.7%	91.4%/1.7%
933	78.4%/3.4%	92.7%/0.3%	84.5%/3.1%	94.9%/0.0%
931,933	83.1%/7.1%	87.7%/2.7%	88.2%/7.0%	90.2%/2.4%
931	84.7%/0.8%	97.1%/0.3%	90.4%/0.3%	90.2%/1.7%

Figure 5-7: Vehicle Headway Percentages excerpt for 4/1/02

their scheduled headway. This standard is neither more stringent nor more lenient than the previous standard, it simply measures a different value. If all vehicles in the system uniformly arrived at 1.49 times the scheduled headway, this figure would show 100.0%/0.0%, while the passenger waiting figure would show 66.6%/0.0% because one third of the passengers, on average, would arrive within the first 0.49 of a scheduled headway and would have to wait more than one headway. A fundamental difference is that a long headway will only effect the fraction of passengers who arrived soon after the preceding train, but some passengers will continue to receive acceptable service. With the vehicle-based measure either the vehicle meets or exceeds the criteria. Also, with this measure, each station is weighted evenly, and each part of the time period is given the same weight. In the passenger-based measures, problems which occur on the shoulders of peaks, for example, are acknowledged to be less serious.

This email can also be quite useful. I found that when a Short Turn was identified, it was guaranteed to be an interesting day to look at the graphics. The user of this summary service can determine if the passenger, vehicle or exception figures also provided warrant further investigation with a visit to a password protected website that makes available for download more detailed text of these figures, showing the passenger waiting or vehicle headway adherence split by station, as well as graphical products which will be described next.

Daily Summary Plots

Graphical summaries of the train position data are created and made available for download from a password protected website. A number of different plots are gener-

ated including:

- A package of overview plots, one file for each line, which cover the day from 2am to 2pm, as well as 2pm to 2am a plot per time period and per direction.
- A different graphical output is a single summary file which contains zooms of a three and one half hour window including the AM [6:30AM-10:00AM] and PM [4:30PM-8:00PM] peak periods, in each direction for each line, as well as the AM and PM overview plots for each line and also a very dense plot of both directions, all day one plot for each line.

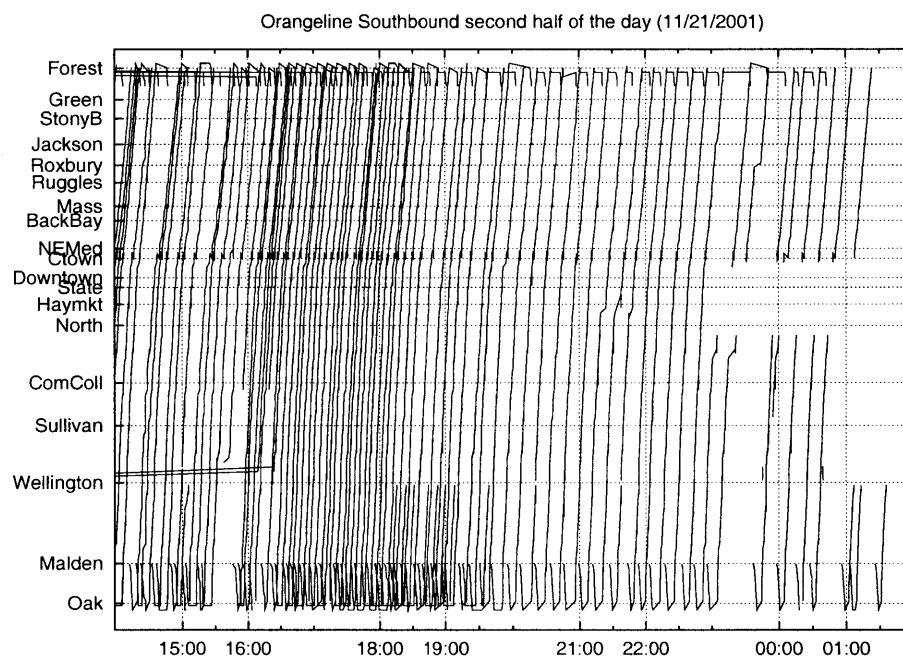


Figure 5-8: Example half day, one direction plot

An example of the detailed time-space diagrams available for each line, for each day is shown in Figure 5-8. The plot shown is an interesting day when many unusual things happened on the line. At the beginning of the plot, on the far left side around 14:10, there is a fifteen minute gap in service with a nominal base headway of eight minutes. This gap in service worsens to 20 minutes as the heavily loaded lead vehicle slows further and further, arriving at Forest Hills around 15:00. The gap in service is not made up for with recovery time at the terminal, nor are other vehicles held in the

Northbound direction to reduce the gap. By the time the slowed vehicle completes the return trip to Oak Grove around 15:50, no vehicle has left that terminal in a half hour. Management fixes the problem for the rest of the line at 15:45 by short turning Northbound trains between Sullivan and Wellington and fifteen minutes later at Community College. These short turned trains fill the gap perfectly and service returns to normal.

In the middle of this situation, at 15:00, a train pulls out of service into the Wellington yards. This figure does not show whether the train was malfunctioning and had to be pulled out of service, or if it was a scheduled repositioning for the peak. Either way, the gap in service this creates also degrades and the gap widens to fifteen minutes by the time it reaches Forest Hills. There are also delays, sensor failures [discontinuities in the tracks towards the end of the day] and an extraordinarily long headway of 40 minutes after the 23:00 train leaves Oak Grove. This would be a very good day for line managers to dissect, and make sure they are happy with all the decisions made. Interventions did assure that the peak demand was well served – headways between 16:00 and 18:00 are generally consistent. However there were long gaps in service at off-peak times which could lead to passenger complaints.

Terminal Operations Visualization

Another use of the data processed by this tool is for the study of terminal operations, an example of which is shown in Figure 5-9. In this figure, the lines indicate train position as estimated from segment occupancy data, with horizontal segments denoting lay-overs. The plus symbol indicates the manually observed bell ring time, the star indicates the observed door-open arrival time, and the box indicates the door close time. The x symbol indicates the system log of the bell time, which is the only automatically collected information which does not agree well with the manually collected data.

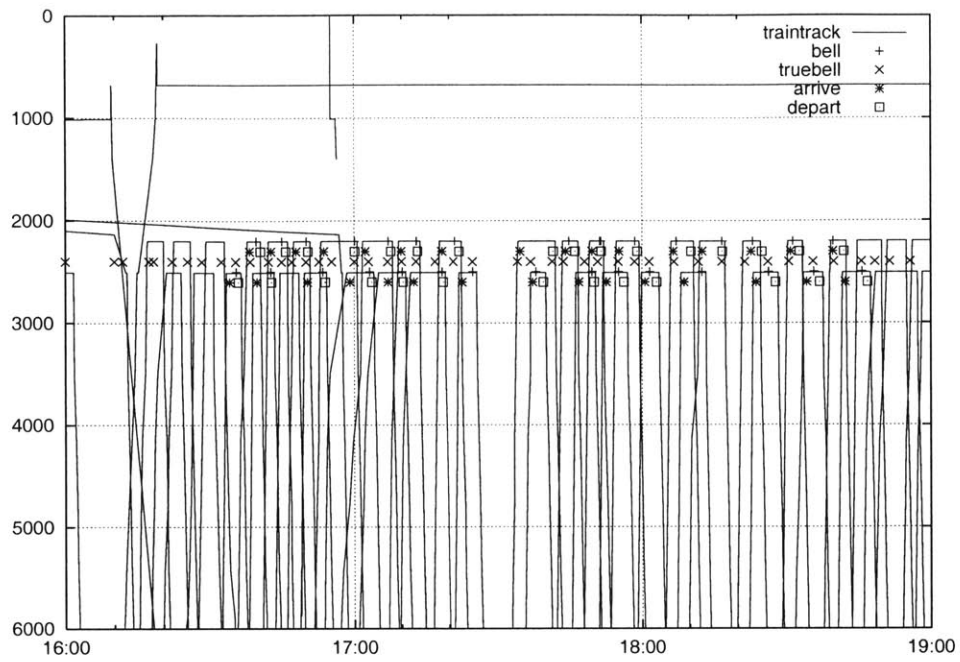


Figure 5-9: Train trajectory and manually collected confirmation.

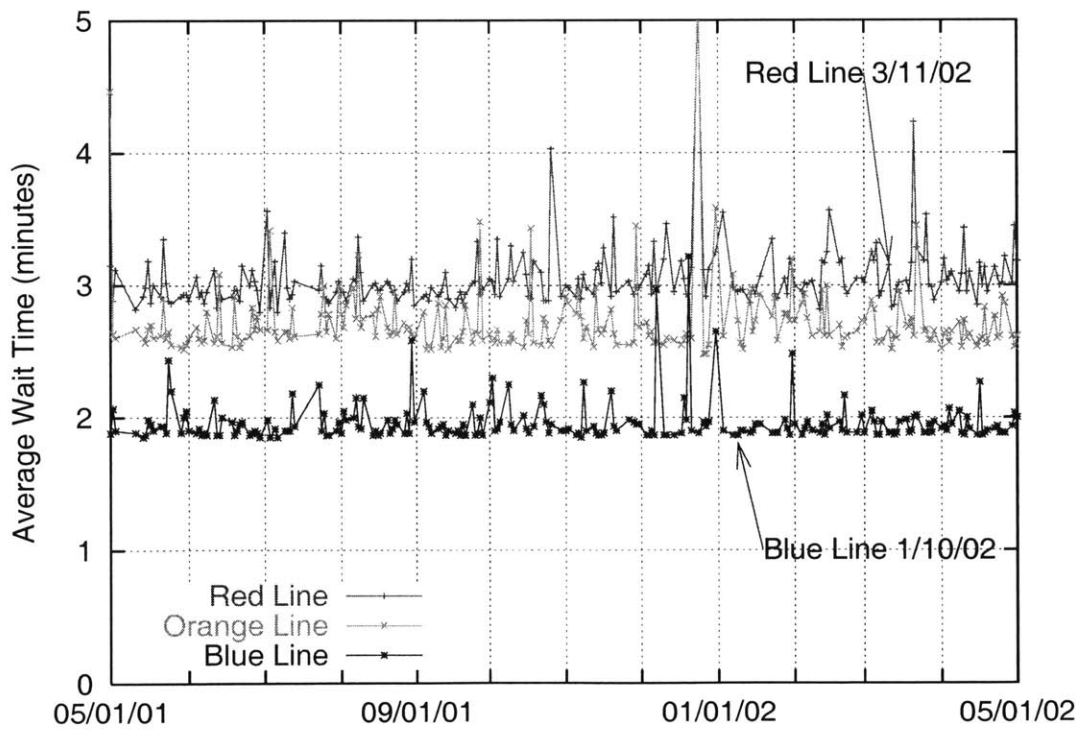


Figure 5-10: Average Passenger Wait Times (minutes) in the AM Peak.

	Blue	Orange	Red
110% minimum observed wait time	2:03	2:45	3:06
Time Period (#weekdays)			
May-August (75)	11	23	14
September-December (74)	13	18	24
January-April (75)	6	28	29
Full Year (224)	30	69	67

Table 5.4: AM Peak Periods With Avg Waiting Exceeding 110% of Minimum

Multi-day Summary Plots

Figure 5-10 is complicated, but provides a lot of information to the user. This figure shows the average passenger wait time, in minutes, for each of the three MBTA heavy rail lines during AM peak on each weekday. This wait time figure is the same as that shown in Figure 5-5, except converted to minutes. It is worth noting that the average value presented for the Red line is a weighted average of the expected passenger waiting times for users of both the branch and trunk portions of the line. This view of transit system operations can facilitate quick comparisons between lines, such as the fact that the Blue line performs very reliably during the AM peak. As shown in Table 5.4 the Blue line has only 30 days with an expected waiting time in the AM Peak 10% more than the minimum observed. This is less than half the comparable number for the other MBTA lines. This figure also makes it apparent that Red line operations in May, June, July and August of 2001 were measurably more consistent than operations in the remainder of the year.

An example of the general high quality of Blue line service is shown in Figure 5-11. The headways, as shown by the distances between parallel lines, are very regular during the peak, with only slight variation from the nominal 2 minute headway. Three extra trains enter the line over the course of the AM peak just before the Orient Heights stop, causing a slight irregularity in headways, but they even out before Maverick. Some slight bunching effect is visible by the end of the line at Bowdoin, but there is not much southbound demand after the Airport station in the AM peak. The high reliability of Blue line service can be explained in part by the fact that it is a short route with relatively low demand.

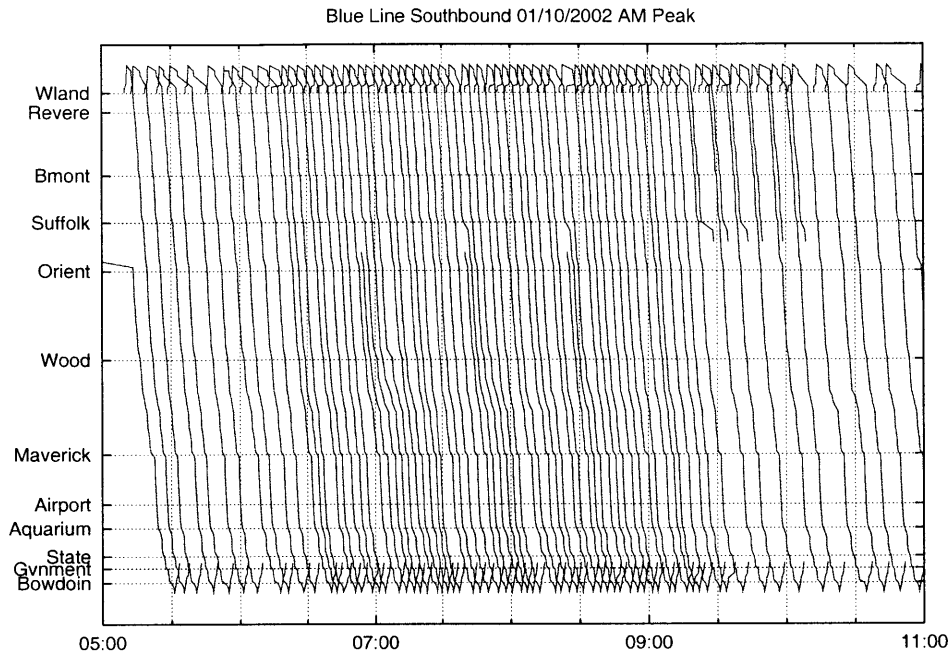


Figure 5-11: 1/10/02 Blue Line Example of Very Regular Service

The Orange line has good periods in the AM peak, with month-long stretches without a major degradation in service quality in this time period. However it does have more than twice the number of problem days as the Blue line. The Red line also often has expected passenger waits well beyond what would be expected assuming reliable service at the current headways. Each of the 67 peak periods with passenger waiting greater than 110% of the minimum are likely to be interesting cases, similar to Figure 5-8. Even a fairly minor deviation from ideal Red line waiting time such as that indicated by the arrow in Figure 5-10 pointing to the average passenger waiting for 3 minutes, 9 seconds on 3/11/02 proves interesting. On this day, shown in Figure 5-12, there were two separate instances of peak period headways over ten minutes in the downtown transfer areas, when the nominal peak period headway is four minutes. This is a deviation that deserves the attention of the controller and line manager to identify why the Ashmont and Braintree trains interleaved to provide a gap in service at 7:30am, as well as what caused the seven minute 7:55am delay before South Station.

Days which appear in the average wait time plot as a major deviation from the

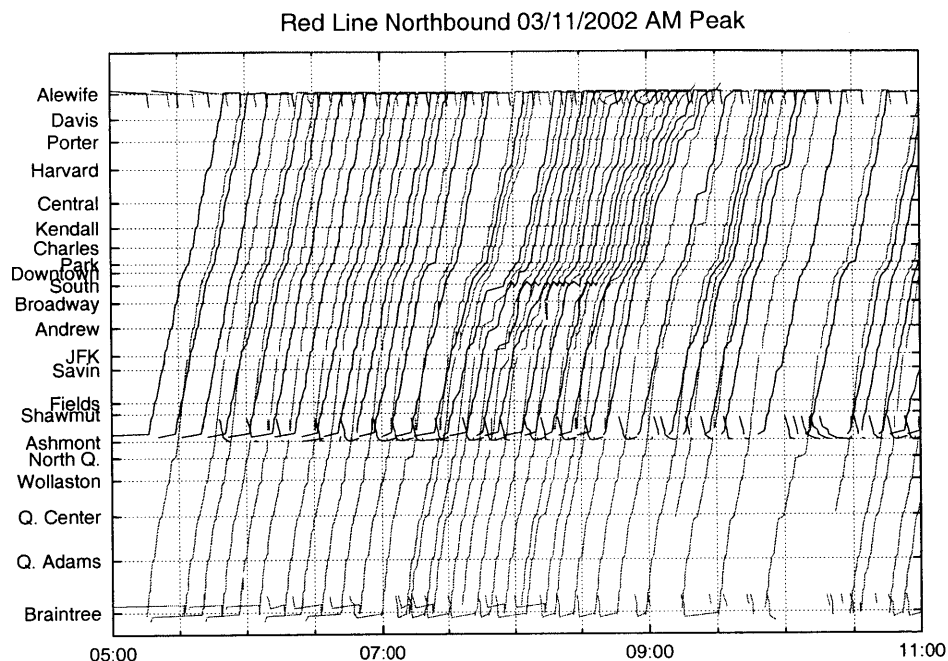


Figure 5-12: 3/11/02 Red Line Example of Mildly Degraded Service

	Blue	Orange	Red
110% minimum observed wait time	4:58	5:57	4:10
Time Period (#weekdays)			
May-August (75)	24	21	43
September-December (74)	7	6	39
January-April (75)	12	27	38
Full Year (224)	43	54	120

Table 5.5: Evening Periods With Avg Waiting Exceeding 110% of Minimum

ideal, such as the 4:20 average wait peak for the Red line on 3/21 are also interesting to study, but typically contain major failures which are harder to accommodate. On 3/21 a train was stopped for more than half an hour starting at 6:20AM at Charles station, which is right after the Downtown portion of the line. This major disruption had ripple effects later in the peak period, which resulted in ten minute headways.

Figure 5-13 shows the same information as that in Figure 5-10, except for the evening hours – between 18:30 and 23:30. This also shows several interesting things. First, to explain the most eye-catching effect, it is comforting to see high quality service for revelers late night on 12/31/01. Unfortunately, as summarized in Table 5.5, typical expected wait times are substantially higher than for the peak, and the vari-

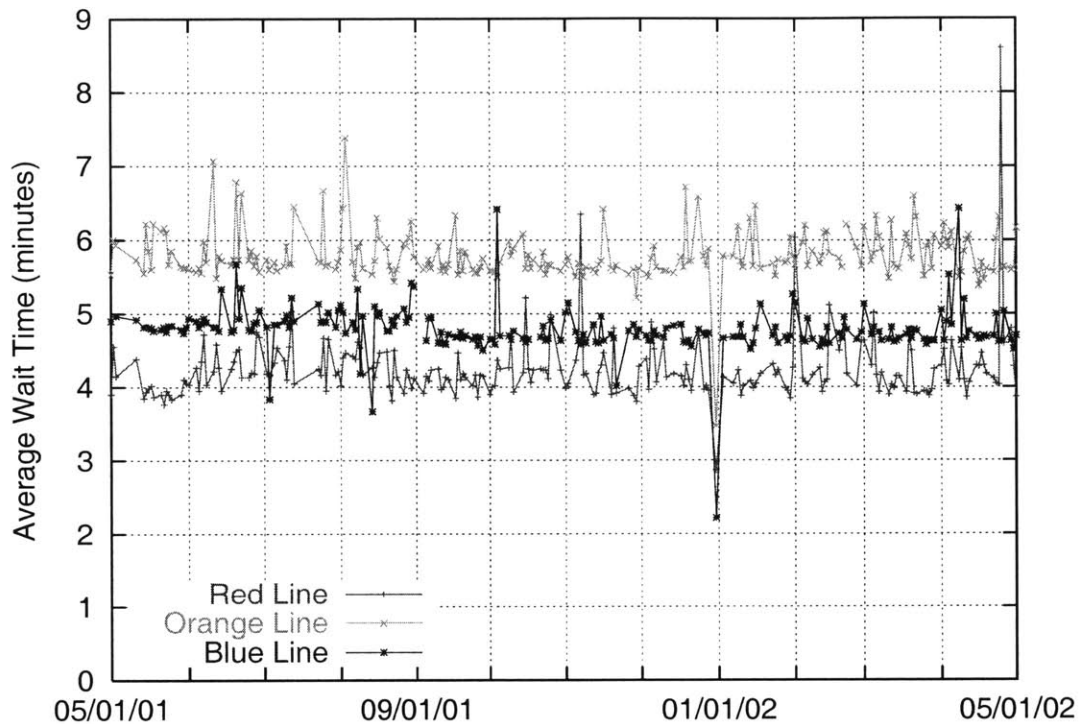


Figure 5-13: Average Passenger Wait Times (minutes) in Evening.

ability is much greater for all the lines. From this view of operations the Orange line in particular appears to be having a great deal of trouble providing consistent service, although it was generally better between September and December of 2001. The Red line is providing on average the least passenger waiting time in this time period, but it is providing the least consistent service of all the lines, by a large margin. More than half the days observed had average evening waiting times greater than 4 minutes 10 seconds; this despite a number of days with evening performance well under 4 minutes, and minimum waiting times under 3 minutes 50 seconds achieved several times in each period.

Information on Processing Requirements

This application provides a great deal of information about processing and storage requirements for a future, upgraded, MBTA performance auditing system implementation at the agency.

Daily position dumps for each of the three heavy rail lines provide the bulk of the

input data. There are approximately 110,000 records total for a given weekday, which require approximately 7.3Mb of storage. Each record gives the segment identifier, the time of segment occupancy, the identity of the the lead car and the scheduled trip identifier. Auxiliary data files available are the consist information “train sheets” from each yard, as well as the status of the terminal start light which signals vehicle departure. These extra files provide approximately 2000 more records, or 0.2Mb of information. There is a fair amount of redundancy in the file format generated from the database dump. These files compress to approximately 10% of their original size, simplifying data transmittal. Given the insensitivity to delay in this post-processing system, the input data could even be transferred over a telephone modem link in less than three minutes.

Initial work to process the data using a spreadsheet based program showed processing times of less than 30 minutes on Pentium II class desktop PC hardware; purpose written software in this application were much faster, generating individual time-space plots in under ten seconds, and updating summary images in a similar time frame. However, ten seconds is too long to wait for a web request to begin completion so the images should be created ahead of time and saved as files, rather than being generated on demand. These plots can be quite sizeable, with a full day summary set of images requiring 2.5Mb of storage. A greater concern is the storage of the processed data products for further research, which take approximately 7.5Mb of space each. The storage requirements for the output data products are the most constrained, at over 10Mb per day. The initial system had only 2Gb of space for system use. Only a half year of data were available at one time, leaving only the full year in compressed products, and small output products such as summary emails. A larger, 40Gb disk was later acquired and now all data is again accessible, requiring only 3.2 Gb for the entire year of uncompressed data files and plots. More recent computer hardware may be able to generate this data dynamically with adequate responsiveness, though generation of plots in advance using obsolete desktop hardware upgraded with an inexpensive high-capacity hard disk has potential to keep hardware costs for such a system extremely modest.

5.2.2 Summary

Overall, the use of this system provides a number of benefits to the agency. The improved ability of rail managers and controllers to visualize and track line performance over time can be an important tool in the provision of high-quality rail transit service.

5.3 Implementation Barriers and Solutions

This section presents solutions to barriers to implementation encountered in the implementation of the MBTA Performance Auditing application. Specifically, I encountered a number of data items available in formats designed only for human legibility. The solution to these was to provide a set of accessible, machine-usable data items. Another issue encountered was difficult to access systems, to combat this I provided a separate data clearinghouse. Lastly I overcame a short time-frame for implementation and lack of database expertise by using standard, freely available gnu / Linux-based tools.

5.3.1 Translate to Machine-Usable Formats

It is desirable that even data-intensive reports are formatted in such a way as to provide a visually appealing hard copy for human perusal. While it was expected that the PDF version of the 1995 15-minute passenger on/off counts was formatted to print well, it was unfortunate that the spreadsheet version was similarly formatted, with data items spread out on the spreadsheet in such a way as to restrict the ability to export from this file for further processing. This required scripting to bring the appropriate columns from the spreadsheet into a simple grid of one row per time period and one column per station. Also, the destination preference of arriving passenger was needed, but this is not part of the available passenger demand data. In order to deal with this lack of detail in the data, it was estimated based on observed loads just after the line split. To calculate the rate of passenger arrival with a branch preference, three figures are used: the fraction of passengers who alight before the

split, denoted C , the fraction who alight after the split on branch A, denoted A , and the fraction who alight after the split on branch B, denoted B . The fraction of passengers traveling solely on the trunk portion, alighting on A and alighting on B was estimated to be $\frac{C}{A+B+C}$, $\frac{A}{A+B+C}$, and $\frac{B}{A+B+C}$ respectively. The fraction is likely to be slightly different, given employment and housing densities, but this estimate was straightforward and deemed adequate.

The extensive ATC data from the MBTA also required automatic reformatting to be usable in processing. The first step in processing the data files is the translation of track segment identifier to position along a route. This is done via a simple table lookup. The next step involves tracking trains through terminals and is necessary because only the lead car in the consist is used to identify a train in the database dump provided. Only one of the seven heavy rail terminals has a loop turnaround; in all the remaining terminals the driver changes ends of the train, and the lead car also changes. This processing step adds a constant identifier to each consist using the consist information from each yard to provide a unique identifier for each consist that makes trips during the day. This facilitates the use of the data for the study of terminal operations, as was shown in the previous section.

Data from the public schedule is also used in this processing. As in the CTA, the schedule is only readily available in a human-readable format – although a different format from the CTA. Fortunately in this case headway based measures are favored, so comparison of vehicles to the schedule is not needed. The only schedule information required is the average headway per line, per time period, which was manually converted.

5.3.2 Build off standard, flexible, data processing tools

The use of freely available, flexible tools to automate data plot generation allowed rapid system implementation and simplified changes to meet new requirements. The reformatting of data products was for processing using standard tools available with the gnu / Linux operating system, including “bash”, “perl” and “gnuplot”. These programs provide the ability to run a sequence of commands simply – the scripting

language “bash”, to reformat data easily – the extraction language “perl”, and to automatically generate graphics given a specification using a flexible tool known as “gnuplot”.

5.3.3 Provide Data Clearinghouse

Another key lesson was in the value of ease of access to information, and the value of having this information available at a range of levels of detail. Train position data has been available from the MBTA ATC system for several years, but it remained inaccessible except in the real-time displays for service management, and in listings of a few records for a vehicle which is the focus of an investigation. Running queries on the database to provide information about a particular days’ operations was difficult due to access restrictions as well as the need for familiarity with the agency database. Automated export of data from this database to a non-mission critical machine was an effective solution to the access problems. This shifting of risk away from MBTA servers was in this case even more extreme than usual as the data was exported from the MBTA to a server at MIT for processing and redistribution. This avoided difficulties in providing external access to agency data.

Daily Summary Email

The automatically generated email format can be quite useful for identifying days with interesting interventions. One opportunity not fully taken advantage of in this system is to use the fact that even the busiest person reads the subjects of all his or her emails. The subject line is an opportunity to make the recipient want to take time to read the contents. A lesson learned is to put the most important couple of numbers from the daily summary in the subject line [for example, total stations experiencing service of greater than 2x the scheduled headway, and the name of the line with the worst performance] both to grab attention and to allow the email subject history to serve as an index.

Data Web Site

The center of the data clearinghouse is the website, which provides a single place to browse for data items of interest. The importance of tailoring information to its target audience was emphasized by the lack of agency interest in the output from this tool. A number of lessons from this are to use cross-platform widely available formats and to provide data at different scales, both in terms of time and in terms of detail. Ease of navigation and quality of visualization are important contributors to usability; the clearinghouse with a simple web interface and multiple levels of detail provides both.

The plots are only available in Postscript format, which is a common printer language among Macintosh users, and is a common display language among Unix and Linux users, but is fairly difficult for the target audience which contains a majority of Microsoft Windows users. Opening the file requires finding and downloading a non standard viewer. A cross-platform standard format such as PDF or PNG would be much easier for distribution, rather than Postscript which is standard only on some operating systems.

Another lesson is to provide not only different levels of detail summary, but also to provide different levels of data resolution. Higher resolution formats slow use and in some cases reduce the output usability by creating large files. Multi-page reports generated by this system are 500kbytes to 2.5mbytes which can take a noticeable time to download. It is important to provide small, quick to load graphical summaries separate from the highly detailed information designed for print out. High resolution data files for use on laser printers provide a wealth of data when needed, but often a quick-loading lower resolution image would suffice to show operations were nominal. Also, having such data readily available is not helpful if the data is not indexed in a way that allows export. Meta summaries available on the website to further direct users to interesting data are a solution to the general agency problem of having so much information that it is no longer useful.

Before this automated processing system became available the only options for

investigating operations on a previous day were limited to quick comments from the memory of the controller, a time-consuming playback of archived data or slow-spreadsheet based tools using a specialized dump of database information. For the most part, only days which were anecdotally known to contain extremely interesting events were reviewed. While the system is not perfect in terms of email and web forms, it is substantially more convenient for quickly skimming a days operations than existed previously.

Chapter 6

Summary and Conclusions

There is potential for significant subsidiary benefits for an integrated approach to the use of automatically collected data within an agency. Automated Vehicle Location systems, Automated Fare Collection systems, and other agency hardware installed to fill particular needs should be viewed as sources of information on vehicles and demand – the state of the transit system. A range of agency functional needs can be met more effectively through monitoring of this digital representation of system state in various ways and on a range of time-scales.

There is a long history of the research use of automatically collected data to address agency needs such as Operations Analysis [33], Operations Monitoring and Passenger Information [21] and Service Management [2] just to name a few. This thesis distinguishes itself by developing a framework exploring a range of agency needs, identifying particularly rewarding applications, and then providing prototypes of those applications.

The range of input sources considered include hardware systems such as for Security / Emergency Location, Automated Vehicle Location, Automatic Vehicle Identification, Automated Train Supervision and Control, Trip Time Analyzers, Automated Passenger Counters, Stop Announcement Systems, and Electronic Fareboxes, among others. Agency needs for which these systems can provide an improved level of performance include Security / Safety, Service Management, Passenger Information, Maintenance, Marketing, Operations Analysis, Performance Monitoring and Audit-

ing, Scheduling, Route Planning, and Budgeting.

Each application identified is reviewed in order to identify synergies between sets of applications using similar intermediate data products, such as segment-level run times and variances of vehicles on a route. Then the relative benefits of each application are compared, as well as the relative difficulty and expected time frames for completion. This ranking of project difficulty will change as more is known about the ease of access to information for each application, when new communications links are added, and when agency computer support staff change. This ranking facilitates managerial involvement and allows the high level decisions to be made to assure applications using automatically collected data have the support to be implemented.

Applications with a high benefit to difficulty ratio which were feasible in a short time-frame were selected for implementation. Implementation lead to a number of agency lessons which are reviewed in the next section. This chapter finishes with a section on future work split into promising case-study agency specific applications, followed by more general areas worthy of investigation in the area of integrated agency data use strategies.

6.1 Conclusions

A number of conclusions can be drawn from the framework application to the CTA Bus and MBTA Rail case studies. Specifically, that the framework outlined in Chapter 3 was productive, there are indeed significant benefits to be captured from further use of extensive data, that having a data repository facilitates this further use of data, that agency use of software packages has hidden costs and the specific requirements for data completeness and update rate needed for use in agency functions.

Many agencies do not collect as much data as they could about their system state. The data that is collected is often not archived for future use, and this collected data is underutilized agency wide. Management needs to have a better understanding of the utility of agency data collection systems for comprehending and improving system performance. Management also needs the information to guide the future

automatic data collection system implementation and use, rather than delegating that responsibility to agency staff who have the technical expertise but may not have a broad understanding of agency goals.

6.1.1 Productive Framework

The framework developed in Chapter 3 was a useful method for organizing an agency's use of automatically collected data. In the framework application to the CTA and MBTA a number of useful projects were identified including the three with prototype implementations.

In order to apply the framework to a new agency, or to a different division of the CTA or MBTA, the two case studies presented here can be used as templates. The first section on review of agency data sources is straightforward – what systems can provide data in electronic format about system vehicles and potential and actual system users. Next the identification of individual applications comes from an appraisal of the extent to which each agency need is currently met, then the specification of the improvement possible using automatically collected data. Section 2.1 can be consulted as a guide, and examples of particular applications used from this section of each case study as described in Section 4.1.3, and Section 5.1.3.

6.1.2 Significant Benefits of Using Data

Each of the case studies have shown that there are significant agency benefits available through the use of automatically collected data. Illustrated in Section 4.2 are benefits to the CTA of more intensive utilization of the Bus Emergency Communications System, including the ability to manage operations, plan more effectively, provide real-time passenger information and observe operator performance. Section 5.2 presents benefits of using data from their Rail Automatic Train Control system to audit overall line performance using passenger-centric measures of performance.

6.1.3 Data Repository

In both cases a large portion of the challenge of the case study was in the manipulation and redistribution of data. Once data is in a standard, comparable format, the incremental difficulty of further applications is much less, as shown in the CTA case study. In order to reduce the barriers to in-house data utilization, agencies should strive for a centralized data repository with the following features:

- Real-time displays should use data from the data repository to decouple display and collection hardware. This implies that data must be available in real-time from the system.
- Data should be mirrored on both production and development servers, to allow real-time application testing with negligible risk to critical applications.
- The repository should present itself as a single, coordinated system to ensure errors detected and remedied by one data user do not adversely impact other users.
- The repository should be as comprehensive as possible, containing dynamic and static information.
- Derivative information should also be redistributed in this centralized repository to reduce duplication of effort.
- Items should be available in a standardized format to enable ease of coordinated use through standard mechanisms for automated comparisons.
- A data history should be available on-line and all prior data should be archived.
- The data should reside in a database as it facilitates coordination – the joining – of different items, but it should be normalized to remove ambiguity and exported to flat files for ease of use in non database-aware tools.

Decouple collection and display

Decoupling the data collection from the data presentation is an important step in isolating the system from the particular hardware used to collect data. A system with such an abstraction will accommodate new sources of data more readily, and will allow tool migration when the underlying data collection hardware inevitably changes. In the case of the MBTA, system performance auditing tools were abandoned when the switch was made to an Automatic Transit Control system.

In order for real-time displays to use data from the data repository, the repository should introduce a minimum of delay between data collection and availability in the repository. This real-time data availability is feasible, as was shown in the CTA real-time auditing and passenger information case-study application.

Repository mirrored on non-critical server

To facilitate real-time application testing, there should be a secondary server which provides access to possibly slightly delayed real-time data. If this secondary server is overloaded by misbehaving programs this should not impact the primary server or the operations the primary server supports. This will simplify the development process for real-time visualization and processing of agency data, and encourage new experimental applications. This sort of arrangement allows testing of applications in a realistic environment, for example with realistic rush-hour data rates.

Coordinated system

The repository should present a unified interface to potential users. While it will likely be composed of several machines, such as primary and development servers, these systems should act in a coordinated way, automatically maintaining a consistent digital representation of the state of the transit system. This integrated data clearinghouse will improve the accuracy of information received.

In order to ensure that the corrections to errors detected and remedied by one data user are propagated to other users, it is important to have one system to which

everyone refers as “truth”. This co-exists well with the above benefits of allowing independence from collection mechanisms, as well as the further benefit of having multiple data users audit the data quality. If a vehicle positioning system is used for performance monitoring and a sensor is malfunctioning and reporting implausible positions, this will be flagged as a maintenance issue rather than having it go undetected until the vehicle position is needed to respond to a silent alarm event triggered on board.

A positive example of how this is starting to take place in agencies relates to the route map component which is now typically held in a Geographic Information System. Before GIS was commonplace, users in operations would manually create maps for operator reference separate from the street and stop level route maps kept by planning. Multiple copies of the same information – route map – were maintained all requiring simultaneous change when a bus was rerouted. A similar effort is needed for other types of agency data.

Comprehensive repository

By incorporating a range of sources of dynamic and static information, the data repository will reduce search time. In both case studies the time spent searching for needed information was substantial, but primarily in the CTA bus case. The CTA is beginning to develop an integrated data repository, but a number of data items were not available in the same location as other information. For example, vehicle position, operator identification and schedule information were all stored in different locations. A large challenge of the CTA application was in making the vehicle position information available in a way that could be easily used. By developing an integrated clearinghouse, future projects will be simplified by reducing the time spent finding information.

A flexible system for managing transit data should have a number of items available from it: vehicle route map and driver / run information, vehicle position-on-route, scheduled vehicle position, as well as passenger load information from Automated Fare Collection systems, Automated Passenger Counters and/or historical

data.

Derivative information

Derivative information should also be redistributed in this centralized repository to reduce duplication of effort and encourage further derivation of more complicated summaries of system status. For example, derived data items of interest to operations, such as terminal departure time per trip may also be of interest to planning. The effort to derive such information could be split. One organizational challenge is that it is easier to create derived data products in a one-off fashion on a desktop computer in a spreadsheet than it is to create a data product that is automatically available in a central location.

One mechanism for splitting this effort is to manually provide a batch of a particular derived data product, as well as the information on how it was derived; any future users could invest the effort in making that particular type of derived information calculated routinely. It is important to at least share algorithms for derivation of each figure in a convenient form, be it compiled program, spreadsheet, or interpreted script.

The availability of derived data was found to be useful in the MBTA case where the derived passenger waiting times were automatically calculated. A further derivative of this – the average and minimum wait times, were only made possible by availability of the waiting times for individual time periods.

Format standardization

Items should be available in a standardized format to facilitate coordinated use. For new systems, the particular standard formats an agency has chosen can be specified as a standard system output, however in other cases this may require a non-disclosure agreement with the vendor of a system. In still other cases this may simply not be possible due to concerns on the part of the vendor about protection of their Intellectual Property. In such cases, it may be possible to purchase an export function from the vendor which can be set up to provide data in a known format. Data should

be reformatted to conform to the protocols in the NTCIP Guide. Refer back to Section 2.5 for more information on the importance of interface standardization. In other cases the vendor may be reluctant to cooperate, in a move to discourage or shut out potential competition in the analysis of system data. Another concern is that the data be available in a format which is readily machine usable, such as a spreadsheet or flat file. Sometimes even when data is available electronically, it is formatted in a manner which is designed for ease of human interpretation, but may be parsed by a program only with difficulty.

There are a range of common differences in data formats, for example in the CTA bus case, in the schedules run numbers are stored as three digits for a given route with the origin garage denoted separately by an alphanumeric character, while in the database a numeral that maps to a garage name is used as a prefix to the run number, making the composite garage-run number unique for a given day. More subtle differences should also be designed to be automatically accommodated to encourage further data use, for example facilitating coordinated use of CTA's BECS and AFC data by automatically identifying and correcting for clock offsets between particular AVL and AFC systems.

Historical archive

Having access to a historical record of agency performance information over time is crucial for a number of agency tasks which fall under the broad category of planning. Data should be archived for analysis since archival is inexpensive; with the quantities of data generated even by entire fleets of train supervision and AVL data should generate at the very upper end say 100 gigabytes of data per year, which can be safely stored on a pair of 120GB EIDE hard drives costing \$300 total.

Most agencies will generate far less data, with a years' worth of MBTA high quality heavy rail train tracking data, along with intermediate data products such as position plots and average waiting times, taking less than 6GB of storage without resorting to compression. This can easily fit on the local hard drive of any new PC. Tasks which integrate large quantities of this data will be increasingly common as

hard drives and computing power become more and more affordable. The operations analysis and performance auditing in the CTA and MBTA case studies, respectively, were made possible thanks to extensive data recording.

Database and flat-files

The greatest benefit to storage of agency data in a database is the ease of quick combinations of related information, especially when historical data is kept and derived data products are integrated into the database. For example, a database with a table containing run number, vehicle run time and time of day can be combined with a separate table on typical run-time for a given time of day to create a difference-from-typical run-time result table. This result table can be combined with a separate table containing run numbers, and driver identifier, providing a result table of difference from typical run-time for each run of each driver. This three line query could be used instead of the special-purpose code written to generate Figure 4-7. Given some training in database use, users within the agency who have more agency-specific information can use this data set as a tool for investigating operations.

One problem is that good DataBase Administrators [DBAs] are currently one of the most highly paid IT specialists, and agencies are understandably reluctant to invest in such expensive staff. However if the database is not properly configured, the storage of agency data in a database can be more of a liability than a benefit. Tables in the database should be normalized, with redundant data removed in internal tables and added back through combination with other tables when reports are generated. Databases should also be indexed for fast queries, especially in databases used in real-time systems. Further, using an understanding of the underlying operations taking place, a database administrator can reorganize queries in a more efficient manner to improve performance. Query tuning and judicious use of database indices in more than one instance led to a reduction in computing requirements by two orders of magnitude, as experienced in the development of the CTA Operations Management and Passenger Information tool. A query to present the most recent 30 minutes of position data for each route originally took approximately 100 seconds of execution

time, which was reduced to 2 seconds of execution time through tuning. These requirements are in addition to the more mundane tasks of assuring adequate disk space and adding users.

Also, flat files will never be completely replaced. Data should be automatically exported to flat files for ease of use in non database-aware tools. Tools such as 'perl' and 'gnuplot', used in the case studies require flat files for processing and plotting. Quick spreadsheet-based analysis can be very valuable for gaining insight into operations. These tools should be accommodated through regular machine-readable report generation from the database.

6.1.4 Hidden Costs of Software Packages

As discussed previously, automatically collected data can be used to provide a deeper understanding of system performance than one based on manual observations alone. To some extent this understanding of line performance has been improved by software purchased to provide a range of standard-form summaries about the agency. However, agencies must carefully consider all the costs of purchasing software packages to process data from new data collection hardware. The hidden costs arise in the reduction in agency productivity caused by reliance on support and maintenance contracts. This reliance on outside support allows agencies to avoid hiring data-savvy staff who can solve agency-specific problems using the range of agency data which is available, not just the data products that vendors designed for.

While it is important not to require employees start from the ground up in processing sources of agency data, a total reliance on commercial packages and support contracts to provide agency-specific customizations is equally wasteful. There is a happy medium between these two extremes which involves carefully specifying commercial software packages which interoperate well with the agency data redistribution infrastructure as well as developing an agency staff comfortable in manipulating large, complicated, data sets. These agency analysts should work closely with agency departments, possibly on rotating assignments inside departments, facilitating integrating the output from a number of agency tools in a manner that is most helpful.

Each of the many functions within the agency which can take advantage of automatically collected data will benefit from the digital representation of system state, and the expertise to leverage it.

6.1.5 Data Quality

It is important to recognize that even low quality data – data with gaps and with poor position information – can be helpful in many applications. The CTA case study illustrates that data collected at a range of different update rates is useful. Data polled every 4 minutes is adequate for alerting management to service problems, such as bunching. Data polled every 100 seconds is shown to be adequate for bus time arrival predictions. Data polled every minute begins to be adequate for identifying sources of service unreliability. The highest update frequency data studied – every 4 seconds – is adequate for signal priority as well as accurate schedule setting.

These figures apply even to data sets which are not complete, such as those with frequent missing vehicles. If bunching is apparent in such a data set, it is sure to be a real problem on the street, and the poor reliability should be addressed. Furthermore, the extended periods of time / distance in Figure 4-3 without any vehicle traces could either be bona-fide gaps in service or they could be malfunctioning vehicles. Some vehicles do not respond to polls, and still others that are on the route are not logged in properly and will not be known to be on the route from the system's point of view. Either alternative – long gap in service, or vehicle not operating properly are worthy of investigation and of supervisory attention.

Depending on the passenger information application, incomplete data may be adequate. If the user is being directed to leave home or office at a particular time to catch a bus, then having position data on only a subset of vehicles may be sufficient. For at-stop passenger information, a missing vehicle may be accepted since the signs will be regarded as an upper limit on waiting time, however if passengers decide to walk rather than wait for a bus based on these numbers, having an unexpected bus pass them may be a source of user frustration. Frequent update rate data can still be used for scheduling if it is incomplete, presuming that vehicles which are not providing

data are typical of the system and their omission is not a cause of bias. Complete data sets, such as generated by the MBTA ATC system allow calculation of a range of interesting data products such as average headway and expected passenger wait time, which are not available in systems with entire vehicles missing from the data set.

6.2 Future Work

This section first reviews avenues for future work on applications for each case-study explored in this thesis. Then it discusses broader issues worthy of further investigation in the area of strategies for integrated agency data use.

6.2.1 CTA Applications

The CTA case study showed the value in a number of uses of automatically collected data. Permanently improving the storage of data, providing a wider distribution for passenger information, bus supervisor adoption of automatically collected data for service management, and directed operator retraining are natural extensions of work in this case study. Another category of work is implementation of applications identified as possible soon or in the near future.

Improving the Data Repository

While data sets are becoming more accessible at the CTA, data is still stored in a number of different locations. Schedule data, AFC data, Bus Operator login data and Bus position data and Rail position data are all available in different places and in different formats. The system in this case study integrated Operator data and position data, and translated between formats of different systems, however this integration was not permanent. Further, a history of the data which is currently in the Oracle database, such as Operator login information, should be archived to facilitate future analysis such as in this thesis.

The system for collecting and archiving bus position data and providing it in the agency standard Oracle database generated intermittent [weekly or every two week] Socket Server crashes. This may be caused by the added outbound [position request] message traffic, or the added inbound [vehicle position report] message traffic, or a combination of both overloading the system in a subtle way, or it could be related to the 10 second polling of the internal proprietary vehicle position, or some subtle effect of the increase in network traffic, server load or Oracle database access. Either way, the importance of having position data in an agency standard location justifies an occasional system reset. Now the BECS system has been upgraded to have vehicles return position on their own, without explicit polling; adding these returned positions to the database and archiving will be much simpler than it was before this project.

Improving Passenger Information

This particular system left a great deal to be desired as far as passenger information is concerned. While it was effective for expert transit users, a simpler interface is needed for mass appeal. The Bus Info project currently underway at the CTA involving at-stop countdown signs provided by Nextbus is a natural extension of the use of this data. The agency should resume entering bus position data into the Oracle database and set up automatic reports from this database to provide needed position data for the Bus Info project.

Improving Service Management

While the data collected from the BECS system in the CTA case study has some significant flaws – it only works for 100 buses at a time, it is missing vehicles and provides erratic positions in the downtown area – it provides useful information on system state back to the control center. Routes for control center supervision can be chosen at random or based on knowledge of particular poorly performing routes. At least three routes can be observed given the constraints, and this will help to identify Operators running fast, implausible schedules, and inadequate or excessive recovery time. The very fact that the operators may be observed may lead to higher quality

service, as running ahead of schedule appeared common despite it being explicitly forbidden.

With all BECS vehicles now providing position reports, routes which have a majority of MDT equipped buses can be monitored in this manner, without the limitation of only 100 buses at a time. Major hurdles to this application are the creation of the interface, integration of schedule information to provide automatic schedule adherence figures, potential union backlash against electronic observation of individual operators and also the training of the controllers, or particular supervisors, in use of the new tools.

Directed Operator Retraining

While the agency is likely to be forbidden to take action on any individual operator based on electronic observations of behavior alone, the electronic observations can be valuable for guiding human observation. The operator variability exploration which led to Figure 4-7 indicates that while 80% of operators on route 77 run within one minute of the average for the route, 20% are more than a minute fast or slow; the operators who are consistently fast and consistently slow could each benefit from more human supervision, and perhaps directed retraining.

Future Applications

A wide range of applications should be considered for implementation at the CTA, including all those discussed in Section 4.1.3. The greatest opportunities for use of data involve the use of Automated Voice Annunciation System [AVAS] data to provide a double check on other systems, such as the BECS and the AFCs. When the APCs are available this will also be useful in checking AFC performance as not only dwell time, position and route will be recorded as with AVAS, but also measures of demand. BECS data will not only be improved by more timely identification of malfunctioning equipment, but also the AVAS integrated login will improve the reliability of operator logins, which will immediately improve the coverage of BECS derived data which uses this login information. This side benefit of AVAS data will

make real-time BECS projects such as the real-time passenger information and service management more appealing.

6.2.2 MBTA Applications

The MBTA case study also showed there is a great deal of value in imaginative use of automatically collected data. Further improving the storage and accessibility of data, wider agency adoption of performance auditing as well as a number of applications for future implementation are all important future uses of agency resources. Especially interesting are future extensive real-time uses of vehicle position information and information derived from it, such as headways.

MBTA Data Repository

As with the CTA case there is more work to be done in collection and redistribution of data. One substantial improvement would be to store the full range of data used to create the MBTA performance auditing application in the agency database. This would involve reformatting and normalizing the passenger count data, and would make the data set much more usable. Storing derivative data products from this application would be useful to facilitate applications which utilize them and perhaps trigger actions automatically on instances of high or low quality service. Provision of a secondary / development database server with real-time information can facilitate a range of future uses of real-time position data, and derivative data products.

Agency Adoption of Performance Auditing

Extending the use of the performance measures produced by this system to performance auditing by Managers with influence over Controller decision is the logical next step for the performance auditing application described in this case study. A step in this direction is a better training program, familiarizing managers and controllers with the opportunities available using this sort of information derived from agency data. After that, using the output from the tool to help Controllers learn

from instances of good and bad reactions to disruptions in service, can improve the quality of service experienced by users. This tool is in a good position to document the improvements in performance made by sharing best practices within the agency.

Future Applications

A wide range of applications should be considered for implementation at the MBTA, including all those discussed in Section 5.1.3. The greatest opportunities for use of data involve redistribution of the high quality rail vehicle position data currently collected by the ATC and AVI systems on the heavy and light rail respectively. There are several uses of this data, including distributing this information in new forms to controllers and to supervisors in the field for line management and coordination, also distribution to system users as passenger information to help time their trips, and also to future analysis systems to provide dispatching and intervention decision support in real time. Each of these tasks is simplified by the development of a redistribution tier, or a slave real-time data repository which simplifies creation and testing of tools which use train-tracking data.

Another promising category of data uses involve using data from the forthcoming Automated Fare Collection [AFC] system to provide information on demand, used in combination with measures of vehicle position to estimate system loads.

6.2.3 Future Research

There are a range of promising avenues for future research in this area. One category of future research is the more rigorous treatment of agency questions that can be answered by viewing agency data as a window into system state. An important area of research is to explore issues in identifying when vehicles are in and out of revenue service. Also it is possible to extend this work but with a system-wide approach. Another category is exploring issues involved in application development between agency employees in the open source model. Lastly, investigation of the most appropriate organizational structure to improve agency utilization of data sources is

an option.

Rigorous Treatment of Agency Applications

There are a large number of important questions that can be answered using sources of agency data. Precursors to unreliable service may be identified and avoided through a thorough understanding of the dynamics, possibly through simulation. There are a range of interesting derived data products that can be explored based on this work; for example combining information on which controller is working when along with the distribution of per-period average passenger waiting times over a long period of time may show controller-dependent effects, such as controllers with extremely effective recovery strategies.

Augmenting Data With In/Out of Service Information

Another important category of future work is to explore the various mechanisms for separating vehicles which are in and out of revenue service as they appear in automatically collected data sets. This has very far-reaching application as agency reliance on automatically collected data increases over time.

The bus system, and to a much lesser extent rail systems may have scheduled and unscheduled out-of-revenue service known as deadheading, and both bus and rail has scheduled and unscheduled in-revenue express service which does not service passengers at intermediate stops. Ideally some static information on deadheading and expressing would be available from the schedule to augment analysis of vehicle progress on a route with its expected in or out of revenue state. If dynamic info on expressing and deadheading is available from logs of interventions can this be used in a fully automated way? Alternately, identifying off-route buses or rail vehicles not even experiencing a base dwell time at stations can provide an indication of an express service, but separating this from deadheading is dependent on knowledge of the vehicle origin – did it just come from a depot or a yard? An exploration of the accuracy of these heuristics would have many applications.

These questions are of practical importance for applications such as the ones in

these case studies. Data for real-time passenger information must indicate whether the vehicle will be stopping at a given stop. For performance auditing, the figures returned can be misleading if trains are deadheaded or expressed, as passengers waiting on the intermediate platforms will erroneously be considered served.

Systematic Approach

The case studies in this work have a somewhat artificial bus / rail division. Future work could explore the value added in system-wide approach to the utilization of automatically collected data.

Sharing of Applications Between Agencies

An exploration of the issues involved in shared application development between agencies can provide an interesting counterpoint to the off-the-shelf or contract-written software versus in-house developed software. Following the open-source model, agencies could collaboratively develop a palette of applications to customize and improve. This work can form a basis for this sort of project, but more effort must be spent on making the code flexible and extensible. This could be an important subject for deeper understanding.

Agency Organization Structure

Lastly, what are the most effective organizational structures for improving agency utilization of data sources? How should this new understanding of agency data sources as windows into system state be implemented in agencies, in terms of personnel and hierarchy. This thesis touched on the issue of organizational structure conducive to good data exploitation, suggesting a core group technology and transit-need savvy programming staff who coordinate frequently, but are rotated through a range of agency departments on a regular basis, bringing the advantages of their systems integration knowledge to a range of audiences. What other alternative structures are possible, and which is likely to be the most effective in practice?

Appendix A

Example Database Usage

This appendix presents a selection of the queries and indices used in the distribution of real-time bus position information using the Oracle database at the CTA. All the examples of database queries provided will be in the Structured Query Language [SQL], with Oracle extensions. This appendix is split in to three sections. First listed are queries related to the polling of a subset of vehicles which are of interest. Second it presents the queries and indices needed for the efficient automatic creation of the database table with a history of vehicle positions. Lastly it provides an example use of this database table – a query which generates a tab delimited text file used as input to the plotting program for the route 20 position information website.

A.1 Vehicle Polling

```
SELECT DISTINCT a.avl_vehicle_id, a.run_id, r.route_id
FROM current_logon a, route_run r, routes_of_interest i
WHERE (a.run_id = r.run_id)
      AND (a.avl_vehicle_id = v.avl_vehicle_id)
      AND (r.route_id = i.route_id)
```

Figure A-1: SQL Code Used in Initial Database Access

Shown in Figure A-1 is the kernel of the selection of the appropriate vehicles for polling. It selects all the vehicles which have an operator logon with a run number

that indicates a vehicle is on an of the routes of interest.

```
SELECT DISTINCT a.avl_vehicle_id, a.run_id, r.route_id, i.rank
FROM current_logon a, route_run r,
     routes_of_interest i, current_vehicle v
WHERE ((a.run_id = r.run_id)
      OR
      (r.run_id =
       to_number(v.assigned_garage || substr(a.run_id,2,3))
       AND (a.avl_vehicle_id = v.avl_vehicle_id )
       AND (v.assigned_garage <> substr(a.run_id, 0, 1))))
      AND (r.route_id = i.route_id)
ORDER BY i.rank, r.route_id
```

Figure A-2: Garage Check Oracle SQL Code Fragment

As described in the “Accommodating conventions” sub-heading of Section 4.3.1, the query was made more complicated in actual use in as shown in Figure A-2. The ‘run_id’ table entry is really a concatenation of the three digit run identifier and a one digit garage identifier. This modification splits this number into garage and run portions. It also returns vehicles which are not in their home or “assigned” garage and which have matching three digit run numbers.

A.2 Frequent Database Position Update

This section reviews some key steps in the extraction of position information from the vendor-supplied proprietary system and into the agency standard Oracle database for redistribution. This section also addresses the performance issues encountered, and the indices needed to provide efficient access to this position information.

Figure A-3 is an example of the creation of tables which hold a substantial ¹ quantity of position history and yet can be frequently accessed and updated due to the indices on entry time and vehicle identifier. Most queries narrow the universe of possible answers first based on the vehicle identifier as well as the time field to select records which were entered recently.

¹Thousands of records.

```

CREATE TABLE veh_position_history
( avl_vehicle_id  INTEGER NOT NULL,
  latitude         FLOAT NOT NULL,
  longitude        FLOAT NOT NULL,
  time            INTEGER NOT NULL
);

CREATE INDEX veh_pos_hist\${avl_vehicle_id} ON
veh_position_history(avl_vehicle_id);

CREATE INDEX veh_pos_hist$time ON
veh_position_history(time);

```

Figure A-3: Position History Table and Index Creation

```

1: while true; do
2:   rsh -l $suser '$socketserver' dta -1 > $newdtapos
3:   comm -1 -3 $curdtapos $newdtapos > $dtanewadds
4:   mv $newdtapos $curdtapos
5:   sqlload $user/$pass control=$newposctlfile
6:   sqlplus $user/$pass @ $updatepos
7:   sleep $cycle
8: done

```

Figure A-4: Code Fragment for Position Transfer to CTA Database

The fragment of a Bourne Shell [/bin/sh] program shown in Figure A-4 starts with a loop which not stop unless interrupted. The second line remotely runs the command 'dta -1' as the user '\$suser' on the socket server, putting the results of that command in the file named '\$newdtapos'. 'dta' is the vendor-provided program which returns the last known position of a vehicle, or all vehicles if the vehicle identifier '-1' is passed. The third line finds the newly changed positions since this command was last run. In line 4 the newest position update is made the official current list of positions, to allow line 3 to work properly next time. The fifth line in Figure A-4 loads the data from the '\$dtanewadds' file into a temporary table in the database. The sixth line does some reformatting of those lines, for example turning the UNIX time ² into a database date formatted object, and inserts this information into tables containing position and direction history for the preceding day. This step takes care to change only a minimal number of position and direction history records in order to improve performance.

A.3 Example Position Database Use

Figure A-5 is an example of an advanced use of the database position and direction history tables. Two optimizations occur before the actual query. The first is the creation of a `local_logins` table. The `active_login` table is an alias for a table on the live, primary CTA database server. As this table is being compared with a number of other tables implicitly in the `WHERE` clause of the query, it provides a great speed improvement to make a local copy. Another optimization is the creation of a `current_time_window` table which contains only one value. The `&&NOW_UNIX` item is a macro which calculates the current time in UNIX format, and the value stored is 15 minutes [900 seconds] before that time. By providing this information in a single-entry database table the system only needs convert the current time into this unusual formation once, and not once for each comparison to time in the position

²On UNIX-like machines times often stored as the number of seconds since Midnight GMT 1/1/1970. The vendor-provided 'dta' tool provides a creation date in this format.


```

CREATE TABLE local_logins AS SELECT * FROM active_view;
CREATE TABLE current_time_window (time NUMBER);
INSERT INTO current_time_window VALUES (&&NOW_UNIX - 900);
BREAK ON vehident SKIP 1;
COLUMN veh FORMAT A8;
...
SELECT  TO_CHAR(p.avl_vehicle_id) veh, TO_CHAR(r.route_id),
        TO_CHAR(p.latitude), TO_CHAR(p.longitude),
        p.time, d.latitude_delta, d.longitude_delta,
        p.avl_vehicle_id * DECODE( (d.longitude_delta),
                                ABS(d.longitude_delta), -1, 1) vehident
FROM local_logins a, route_run r, routes_of_interest i,
     veh_position_history p, veh_direction_history d,
     current_time_window t
WHERE
    t.time < p.time
    AND d.time=p.time
    AND a.run_id = r.run_id
    AND r.route_id = i.route_id
    AND p.avl_vehicle_id = a.avl_vehicle_id
    AND d.avl_vehicle_id = p.avl_vehicle_id
ORDER BY rid, veh, p.time

```

Figure A-5: Usage of Database Position History Data

and direction history tables.

The query itself is fairly straightforward, providing records which are less than fifteen minutes old for vehicles on routes in the `routes_of_interest` table. Extra conditions are needed to assure that the direction and position history records match to provide both position and direction in one record.

A number of simple SQL commands similar to 'COLUMN veh FORMAT A8' were omitted and replaced by ... for the sake of space. Note that formatting constraints drive the conversion of numbers to strings in the output, which makes values left justified. The formatting commands are important in this example, as they create a tab delimited file which can be easily used by `gnuplot`, the flexible graphics program used for all graphics creation in these case studies. This graphics program utilizes blank lines in a input data file to place a break in a line. Blank lines are generated whenever the vehicle changes direction by the combination of 'BREAK ON vehident SKIP 1' line which adds a new line when the value changes, as well as the definition of the `vehident` column in the output, which changes value when the vehicle travels in the

reverse direction.

This query is run every minute, and immediately after the resulting tab separated value file is used to automatically plot vehicle trajectories, as shown in the red lines on the bus position history website³.

³Refer to in Figure 4-3.

Bibliography

- [1] Sufiya Abdur-Rahman. CTA opens digital eyes at train stops. *Chicago Tribune*, December 2002.
- [2] David P. Barker. Communication, information and responsibility distribution strategies for effective real-time transit service management. Master's thesis, Massachusetts Institute of Technology, May 2002.
- [3] CTA resident outreach web page, October 2002.
<http://www.yourcta.com/welcome/resident.html>.
- [4] Maria del Pilar Rodriguez. Moving from conventional bus service towards bus rapid transit: Establishing priorities. Master's thesis, Massachusetts Institute of Technology, February 2003.
- [5] John Dobies. Bus route evaluation standards. TCRP synthesis, Transportation Research Board, June 1996.
- [6] Planning Dr. Michael Shiffer, Vice President and Development CTA. Leveraging technology to reshape transportation planning. MIT Center For Transportation and Logistics Speaker Series Presentation Friday, February 21, 2003.
- [7] Xu Jun Eberlein. Real time control strategies in tranist operations: Models and analysis. Master's thesis, Massachusetts Institute of Technology, June 1995.
- [8] Markus Friedrich and Helmut Prungel. open.p - an open database for public transit operators and agencies. *Verkehr und Technik*, October 2000.

- [9] Peter G. Furth. Innovative sampling plans for estimating transit passenger-kilometers. *Transportation Research Record 1618*, 1998.
- [10] Peter G. Furth. Data collection for bus planning and monitoring. TCRP synthesis, Transportation Research Board, September 1999.
- [11] Garmin GPS 16 sensor brochure. <http://www.garmin.com/products/gps16/>.
- [12] HASTUS Transit scheduling, customer information and daily operations, October 2002. http://www.giro.ca/webdoc/hastusa_web.pdf.
- [13] Robert C. Herguth. Many CTA execs seldom ride transit. *Chicago Sun-Times*, December 2002.
- [14] Robert C. Herguth and Gary Wisby. 70 *From Chicago Sun Times*, February 2003. <http://www.suntimes.com/output/news/cst-nws-bus12.html>.
- [15] G. Curtis Herrick, editor. *NTCIP 9001 The NTCIP Guide ... updated version 3*. American Association of State Highway and Transportation Officials (AASHTO), ITE, and NEMA, October 2002. <http://www.ntcip.org/library/documents/pdf/9001v0302b.pdf>.
- [16] Jon Hilkevitch. CTA plan has buses talking all over town. *Chicago Tribune*, October 2002.
- [17] A.X. Horbury. Using non-real-time automatic vehicle location data to improve bus service. *Transportation Research Part B 33*, pages p559–579, 1999.
- [18] Automatic passenger counting with GPS-AVL, October 2002. http://www2.infodev.ca/download/bus_infodev.pdf.
- [19] <http://www.init-ka.de/> October 2002.
- [20] Daron Jones. Case studies in navigation and positioning. *UnderWater Magazine*, Jan/Feb 2002.

- [21] C.F.J. Lam. The application of information technologies to public transportation. Master's thesis, Massachusetts Institute of Technology, 1994.
- [22] Derek Sze-Ming Lee. Understanding capacity and performance of urban rail transit terminals. Master's thesis, Massachusetts Institute of Technology, May 2002.
- [23] *MBTA Service Delivery Policy*, September 1996.
- [24] Pratap Misra. GPS Performance in Navigation March 4, 1998. MIT Lincoln Laboratory 1998 GPS Course notes, 1998.
- [25] Angela Mary Moore. A process for improving transit service management during disruptions. Master's thesis, Massachusetts Institute of Technology, February 2003.
- [26] T. H. J. Muller and P. G. Furth. Trip time analyzers: Key to transit service quality. In *Transportation Research Board Proceedings*, Jan 2001.
- [27] Next stop announcements will make bus rides easier for visually or hearing impaired customers. CTA Press Release 10/3/2002.
- [28] Orbital Sciences Corporation, Germantown, MD. *OrbTrac Message Definition*, September 1998.
- [29] Service Planning. *Chicago Transit Authority Service Standards*, July 2001.
- [30] CTA productivity statistics CTA bus ridership, November 2002. <http://www.yourcta.com/news/newspostdescs/132346nov02bus.pdf>.
- [31] October 2002. <http://www.ptv.de/>.
- [32] Sari Radin, David Jackson, David Rosner, and Sean Peirce. *Trip Planning State of the Practice*. Federal Transit Administration, U.S. Department of Transportation, July 2002.

- [33] Adam Rahbee. Rail transit operations analysis: Framework and applications. Master's thesis, Massachusetts Institute of Technology, September 2001.
- [34] RSM services ridecheck plus software summary, October 2002. http://www.rsmservices.com/rc_summary.htm.
- [35] J.G. Strathman, K.J. Dueker, T.J. Kimpel, R.L Gerhart, K. Turner, D. Griffin, and S. Callas. Bus transit operations control: Review and an experiment involving tri-met's automated bus dispatching system. *Transportation Research Board*, January 2001.
- [36] J.G. Strathman, K.J. Dueker, T.J. Kimpel, R.L Gerhart, K. Turner, P. Taylor, S. Callas, D. Griffin, and J. Hopper. Automated bus dispatching, operations control, and service reliability baseline analysis. *Transportation Research Record 1666*, January 1999.
- [37] Michael S. Townes. Passenger counting technologies and procedures. TCRP synthesis, Transportation Research Board, 1998.
- [38] Commissioner Tristani. *Third Report and Order FCC 99-245*. Federal Communications Commission, September 1999. Concerns e911 performance.
- [39] Do you know where your bus is?, January 1995. Penn State University Press Release.
- [40] N. H. M. Wilson, P. G. Furth, J. Attanucci, and I. Burns. *Transit Data Collection Design Manual*. U.S. Department of Transportation, June 1985.
- [41] N. H. M. Wilson, D. Nelson, A. Palmere, T. H. Grayson, and C. Cederquist. Service-quality monitoring for high-frequency transit lines. *Transportation Research Record 1349*, pages 3–11, 1992.
- [42] Corey J. Wong. Improving the network performance of urban transit systems. Master's thesis, Massachusetts Institute of Technology, September 2000.